

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010327 A2

(51) International Patent Classification ⁷ :		C12Q	US	60/308,901 (CIP)
(21) International Application Number:		PCT/US02/14199	Filed on	31 July 2001 (31.07.2001)
(22) International Filing Date:		2 May 2002 (02.05.2002)	US	60/330,292 (CIP)
(25) Filing Language:		English	Filed on	18 October 2001 (18.10.2001)
(26) Publication Language:		English	US	60/288,063 (CIP)
(30) Priority Data:			Filed on	2 May 2001 (02.05.2001)
60/288,063	2 May 2001 (02.05.2001)	US	US	60/289,087 (CIP)
60/288,395	3 May 2001 (03.05.2001)	US	Filed on	7 May 2001 (07.05.2001)
60/289,087	7 May 2001 (07.05.2001)	US	US	60/289,818 (CIP)
60/289,818	9 May 2001 (09.05.2001)	US	Filed on	9 May 2001 (09.05.2001)
60/289,817	9 May 2001 (09.05.2001)	US	US	60/313,416 (CIP)
60/290,194	11 May 2001 (11.05.2001)	US	Filed on	17 August 2001 (17.08.2001)
60/290,753	14 May 2001 (14.05.2001)	US	US	60/289,817 (CIP)
60/291,181	15 May 2001 (15.05.2001)	US	Filed on	5 September 2001 (05.09.2001)
60/291,243	16 May 2001 (16.05.2001)	US	US	60/290,194 (CIP)
60/292,001	18 May 2001 (18.05.2001)	US	Filed on	11 May 2001 (11.05.2001)
60/292,374	21 May 2001 (21.05.2001)	US	US	60/325,683 (CIP)
60/292,587	22 May 2001 (22.05.2001)	US	Filed on	27 September 2001 (27.09.2001)
60/293,107	23 May 2001 (23.05.2001)	US	US	60/336,909 (CIP)
60/293,747	25 May 2001 (25.05.2001)	US	Filed on	3 December 2001 (03.12.2001)
60/294,110	29 May 2001 (29.05.2001)	US	US	60/290,753 (CIP)
60/294,109	29 May 2001 (29.05.2001)	US	Filed on	14 May 2001 (14.05.2001)
60/294,434	30 May 2001 (30.05.2001)	US	US	60/291,181 (CIP)
60/294,827	31 May 2001 (31.05.2001)	US	Filed on	15 May 2001 (15.05.2001)
60/304,879	12 July 2001 (12.07.2001)	US	US	60/291,243 (CIP)
60/308,901	31 July 2001 (31.07.2001)	US	Filed on	16 May 2001 (16.05.2001)
60/312,270	14 August 2001 (14.08.2001)	US	US	60/292,001 (CIP)
60/313,416	17 August 2001 (17.08.2001)	US	Filed on	18 May 2001 (18.05.2001)
60/318,463	10 September 2001 (10.09.2001)	US	US	60/292,374 (CIP)
60/325,683	27 September 2001 (27.09.2001)	US	Filed on	21 May 2001 (21.05.2001)
60/330,292	18 October 2001 (18.10.2001)	US	US	60/312,270 (CIP)
60/333,873	28 November 2001 (28.11.2001)	US	Filed on	14 August 2001 (14.08.2001)
60/336,909	3 December 2001 (03.12.2001)	US	US	60/292,587 (CIP)
60/337,552	3 December 2001 (03.12.2001)	US	Filed on	22 May 2001 (22.05.2001)
60/359,245	21 February 2002 (21.02.2002)	US	US	60/359,245 (CIP)
10/136,826	1 May 2002 (01.05.2002)	US	Filed on	21 February 2001 (21.02.2001)
(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:			US	60/293,107 (CIP)
			Filed on	23 May 2001 (23.05.2001)
			US	60/294,110 (CIP)
			Filed on	29 May 2001 (29.05.2001)
			US	60/304,879 (CIP)
			Filed on	12 July 2001 (12.07.2001)
			US	60/293,747 (CIP)
			Filed on	25 May 2001 (25.05.2001)

[Continued on next page]

(54) Title: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

(57) Abstract: The present invention provides novel isolated polynucleotides and small molecule target polypeptides encoded by the polynucleotides. Antibodies that immunospecifically bind to a novel small molecule target polypeptide or any derivative, variant, mutant or fragment of that polypeptide, polynucleotide or antibody are disclosed, as are methods in which the small molecule target polypeptide, polynucleotide and antibody are utilized in the detection and treatment of a broad range of pathological states. More specifically, the present invention discloses methods of using recombinantly expressed and/or endogenously expressed proteins in various screening procedures for the purpose of identifying therapeutic antibodies and therapeutic small molecules associated with diseases.



WO 03/010327 A2



US	60/294,434 (CIP)
Filed on	30 May 2001 (30.05.2001)
US	60/318,463 (CIP)
Filed on	10 September 2001 (10.09.2001)
US	60/294,109 (CIP)
Filed on	29 May 2001 (29.05.2001)
US	60/333,873 (CIP)
Filed on	28 November 2001 (28.11.2001)
US	60/337,552 (CIP)
Filed on	3 December 2001 (03.12.2001)
US	60/294,827 (CIP)
Filed on	31 May 2001 (31.05.2001)
US	Not furnished (CIP)
Filed on	1 May 2002 (01.05.2002)

New Haven, CT 06515 (US). **MACDOUGALL, John, R.** [CA/US]; 117 Russell Street, Hamden, CT 06517 (US). **PADIGARU, Muralidhara** [IN/US]; 71 Hampton Park, Branford, CT 06405 (US). **RASTELLI, Luca** [IT/US]; 52 Pepperbush Lane, Guilford, CT 06437 (US). **SHENOY, Suresh, G.** [IN/US]; 15 Millwood Drive, Branford, CT 06405 (US). **GERLACH, Valerie, L.** [US/US]; 18 Rock Pasture Road, Branford, CT 06405 (US). **SHIMKETS, Richard, A.** [US/US]; 5 Indian Meadows Drive, Guilford, CT 06437 (US). **ZHONG, Mei** [CA/US]; 45 Harrison Avenue, Apartment 1B, Branford, CT 06405 (US). **EDINGER, Shlomit, R.** [US/US]; 766 Edgewood Avenue, New Haven, CT 06515 (US). **ELLERMAN, Karen** [US/US]; 87 Montoya Drive, Branford, CT 06405 (US).

(71) **Applicant (for all designated States except US):** **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).

(74) **Agent:** **ELRIFI, Ivor, R.;** Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P. C., One Financial Center, Boston, MA 02111 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants (for US only):** **MILLER, Charles, E.** [US/US]; 98 Saddle Hill Drive, Guilford, CT 06437 (US). **KEKUDA, Ramesh** [IN/US]; 1213 Avalon Valley Drive, Danbury, CT 06810 (US). **MALYANKAR, Uriel, M.** [IN/US]; 229 Branford Road, Number 330, Branford, CT 06405 (US). **LI, Li** [CN/US]; 56 Jerimoth Drive, Branford, CT 06405 (US). **PENA, Carol, E., A.** [US/US]; 604 Orange Street, Number 2, New Haven, CT 06511 (US). **SPYTEK, Kimberly, A.** [US/US]; 28 Court Street, Number 1, New Haven, CT 06511 (US). **GORMAN, Linda** [US/US]; 329 Monticello Drive, Branford, CT 06405 (US). **GUO, Xiaojia** [CN/US]; 713 Robert Frost Drive, Branford, CT 06405 (US). **FERNANDES, Elma, R.** [IN/US]; 77 Florence Road, 2B, Branford, CT 06405 (US). **SMITHSON, Glenna** [US/US]; 125 Michael Drive, Guilford, CT 06405 (US). **STONE, David, J.** [US/US]; 223 Whitehorn Drive, Guilford, CT 06437 (US). **ZERHUSEN, Bryan, D.** [US/US]; 337 Monticello Drive, Branford, CT 06405 (US). **PATTURAJAN, Meera** [IN/US]; 45 Harrison Avenue, Apartment 1C, Branford, CT 06405 (US). **ANDERSON, David, W.** [US/US]; 85 Montoya Drive, Branford, CT 06405 (US). **MEZES, Peter, S.** [CA/US]; 7 Clark's Lane, Old Lyme, CT 06371 (US). **PEYMAN, John, A.** [US/US]; 336 West Rock Avenue,

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Novel Proteins and Nucleic Acids Encoding Same

FIELD OF THE INVENTION

5 The present invention relates to novel polypeptides that are targets of small molecule drugs and that have properties related to stimulation of biochemical or physiological responses in a cell, a tissue, an organ or an organism. More particularly, the novel polypeptides are gene products of novel genes, or are specified biologically active fragments or derivatives thereof. Methods of use encompass diagnostic and prognostic
10 assay procedures as well as methods of treating diverse pathological conditions.

BACKGROUND

Eukaryotic cells are characterized by biochemical and physiological processes which under normal conditions are exquisitely balanced to achieve the preservation and propagation of the cells. When such cells are components of multicellular organisms such as vertebrates, or more particularly organisms such as mammals, the regulation of the biochemical and physiological processes involves intricate signaling pathways. Frequently, such signaling pathways are constituted of extracellular signaling proteins, cellular receptors that bind the signaling proteins and signal transducing components located within the cells.

Signaling proteins may be classified as endocrine effectors, paracrine effectors or autocrine effectors. Endocrine effectors are signaling molecules secreted by a given organ into the circulatory system, which are then transported to a distant target organ or tissue. The target cells include the receptors for the endocrine effector, and when the endocrine effector binds, a signaling cascade is induced. Paracrine effectors involve secreting cells and receptor cells in close proximity to each other, for example two different classes of cells in the same tissue or organ. One class of cells secretes the paracrine effector, which then reaches the second class of cells, for example by diffusion through the extracellular fluid. The second class of cells contains the receptors for the paracrine effector; binding of the effector results in induction of the signaling cascade that elicits the corresponding biochemical or physiological effect. Autocrine effectors are highly analogous to paracrine effectors, except that the same cell type that secretes the autocrine effector also contains the receptor. Thus the autocrine effector binds to receptors on the same cell, or on identical neighboring cells. The binding process then elicits the characteristic biochemical or physiological effect.

Signaling processes may elicit a variety of effects on cells and tissues including by way of nonlimiting example induction of cell or tissue proliferation, suppression of growth or proliferation, induction of differentiation or maturation of a cell or tissue, and suppression of differentiation or maturation of a cell or tissue.

Many pathological conditions involve dysregulation of expression of important effector proteins. In certain classes of pathologies the dysregulation is manifested as diminished or suppressed level of synthesis and secretion protein effectors. In a clinical

setting a subject may be suspected of suffering from a condition brought on by diminished or suppressed levels of a protein effector of interest. Therefore there is a need to be able to assay for the level of the protein effector of interest in a biological sample from such a subject, and to compare the level with that characteristic of a nonpathological condition.

- 5 There further is a need to provide the protein effector as a product of manufacture. Administration of the effector to a subject in need thereof is useful in treatment of the pathological condition, or the protein effector deficiency or suppression may be favorably acted upon by the administration of another small molecule drug product. Accordingly, there is a need for a method of treatment of a pathological condition brought on by a
- 10 diminished or suppressed levels of the protein effector of interest.

Small molecule targets have been implicated in various disease states or pathologies. These targets may be proteins, and particularly enzymatic proteins, which are acted upon by small molecule drugs for the purpose of altering target function and achieving a desired result. Cellular, animal and clinical studies can be performed to

15 elucidate the genetic contribution to the etiology and pathogenesis of conditions in which small molecule targets are implicated in a variety of physiologic, pharmacologic or native states. These studies utilize the core technologies at CuraGen Corporation to look at differential gene expression, protein-protein interactions, large-scale sequencing of expressed genes and the association of genetic variations such as, but not limited to, single

20 nucleotide polymorphisms (SNPs) or splice variants in and between biological samples from experimental and control groups. The goal of such studies is to identify potential avenues for therapeutic intervention in order to prevent, treat the consequences or cure the conditions.

In order to treat diseases, pathologies and other abnormal states or conditions in

25 which a mammalian organism has been diagnosed as being, or as being at risk for becoming, other than in a normal state or condition, it is important to identify new therapeutic agents. Such a procedure includes at least the steps of identifying a target component within an affected tissue or organ, and identifying a candidate therapeutic agent that modulates the functional attributes of the target. The target component may be

30 any biological macromolecule implicated in the disease or pathology. Commonly the target is a polypeptide or protein with specific functional attributes. Other classes of macromolecule may be a nucleic acid, a polysaccharide, a lipid such as a complex lipid or

a glycolipid; in addition a target may be a sub-cellular structure or extra-cellular structure that is comprised of more than one of these classes of macromolecule. Once such a target has been identified, it may be employed in a screening assay in order to identify favorable candidate therapeutic agents from among a large population of substances or compounds.

5 In many cases the objective of such screening assays is to identify small molecule candidates; this is commonly approached by the use of combinatorial methodologies to develop the population of substances to be tested. The implementation of high throughput screening methodologies is advantageous when working with large, combinatorial libraries of compounds.

10

SUMMARY OF THE INVENTION

The invention is based in part upon the discovery of nucleic acid sequences encoding novel polypeptides. These nucleic acids and polypeptides, as well as derivatives, homologs, analogs and fragments thereof, will hereinafter be collectively designated as "NOVX" nucleic acid, which represents the nucleotide sequence selected from the group
15 consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or polypeptide sequences, which represents the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101.

In one aspect, the invention provides an isolated polypeptide comprising a mature form of a NOVX amino acid. One example is a variant of a mature form of a NOVX
20 amino acid sequence, wherein any amino acid in the mature form is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed. The amino acid can be, for example, a NOVX amino acid sequence or a variant of a NOVX amino acid sequence, wherein any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no
25 more than 15% of the amino acid residues in the sequence are so changed. The invention also includes fragments of any of these. In another aspect, the invention also includes an isolated nucleic acid that encodes a NOVX polypeptide, or a fragment, homolog, analog or derivative thereof.

Also included in the invention is a NOVX polypeptide that is a naturally occurring
30 allelic variant of a NOVX sequence. In one embodiment, the allelic variant includes an amino acid sequence that is the translation of a nucleic acid sequence differing by a single

nucleotide from a NOVX nucleic acid sequence. In another embodiment, the NOVX polypeptide is a variant polypeptide described therein, wherein any amino acid specified in the chosen sequence is changed to provide a conservative substitution. In one embodiment, the invention discloses a method for determining the presence or amount of the NOVX polypeptide in a sample. The method involves the steps of: providing a sample; introducing the sample to an antibody that binds immunospecifically to the polypeptide; and determining the presence or amount of antibody bound to the NOVX polypeptide, thereby determining the presence or amount of the NOVX polypeptide in the sample. In another embodiment, the invention provides a method for determining the presence of or predisposition to a disease associated with altered levels of a NOVX polypeptide in a mammalian subject. This method involves the steps of: measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and comparing the amount of the polypeptide in the sample of the first step to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, the disease, wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

In a further embodiment, the invention includes a method of identifying an agent that binds to a NOVX polypeptide. This method involves the steps of: introducing the polypeptide to the agent; and determining whether the agent binds to the polypeptide. In various embodiments, the agent is a cellular receptor or a downstream effector.

In another aspect, the invention provides a method for identifying a potential therapeutic agent for use in treatment of a pathology, wherein the pathology is related to aberrant expression or aberrant physiological interactions of a NOVX polypeptide. The method involves the steps of: providing a cell expressing the NOVX polypeptide and having a property or function ascribable to the polypeptide; contacting the cell with a composition comprising a candidate substance; and determining whether the substance alters the property or function ascribable to the polypeptide; whereby, if an alteration observed in the presence of the substance is not observed when the cell is contacted with a composition devoid of the substance, the substance is identified as a potential therapeutic agent. In another aspect, the invention describes a method for screening for a modulator of activity or of latency or predisposition to a pathology associated with the NOVX

polypeptide. This method involves the following steps: administering a test compound to a test animal at increased risk for a pathology associated with the NOVX polypeptide, wherein the test animal recombinantly expresses the NOVX polypeptide. This method involves the steps of measuring the activity of the NOVX polypeptide in the test animal
5 after administering the compound of step; and comparing the activity of the protein in the test animal with the activity of the NOVX polypeptide in a control animal not administered the polypeptide, wherein a change in the activity of the NOVX polypeptide in the test animal relative to the control animal indicates the test compound is a modulator of latency of, or predisposition to, a pathology associated with the NOVX polypeptide. In
10 one embodiment, the test animal is a recombinant test animal that expresses a test protein transgene or expresses the transgene under the control of a promoter at an increased level relative to a wild-type test animal, and wherein the promoter is not the native gene promoter of the transgene. In another aspect, the invention includes a method for modulating the activity of the NOVX polypeptide, the method comprising introducing a
15 cell sample expressing the NOVX polypeptide with a compound that binds to the polypeptide in an amount sufficient to modulate the activity of the polypeptide.

The invention also includes an isolated nucleic acid that encodes a NOVX polypeptide, or a fragment, homolog, analog or derivative thereof. In a preferred embodiment, the nucleic acid molecule comprises the nucleotide sequence of a naturally
20 occurring allelic nucleic acid variant. In another embodiment, the nucleic acid encodes a variant polypeptide, wherein the variant polypeptide has the polypeptide sequence of a naturally occurring polypeptide variant. In another embodiment, the nucleic acid molecule differs by a single nucleotide from a NOVX nucleic acid sequence. In one embodiment, the NOVX nucleic acid molecule hybridizes under stringent conditions to the nucleotide
25 sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or a complement of the nucleotide sequence. In another aspect, the invention provides a vector or a cell expressing a NOVX nucleotide sequence.

In one embodiment, the invention discloses a method for modulating the activity of a NOVX polypeptide. The method includes the steps of: introducing a cell sample
30 expressing the NOVX polypeptide with a compound that binds to the polypeptide in an amount sufficient to modulate the activity of the polypeptide. In another embodiment, the invention includes an isolated NOVX nucleic acid molecule comprising a nucleic acid

sequence encoding a polypeptide comprising a NOVX amino acid sequence or a variant of a mature form of the NOVX amino acid sequence, wherein any amino acid in the mature form of the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed.

- 5 In another embodiment, the invention includes an amino acid sequence that is a variant of the NOVX amino acid sequence, in which any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed.

- 10 In one embodiment, the invention discloses a NOVX nucleic acid fragment encoding at least a portion of a NOVX polypeptide or any variant of the polypeptide, wherein any amino acid of the chosen sequence is changed to a different amino acid, provided that no more than 10% of the amino acid residues in the sequence are so changed. In another embodiment, the invention includes the complement of any of the NOVX nucleic acid molecules or a naturally occurring allelic nucleic acid variant. In
15 another embodiment, the invention discloses a NOVX nucleic acid molecule that encodes a variant polypeptide, wherein the variant polypeptide has the polypeptide sequence of a naturally occurring polypeptide variant. In another embodiment, the invention discloses a NOVX nucleic acid, wherein the nucleic acid molecule differs by a single nucleotide from a NOVX nucleic acid sequence.

- 20 In another aspect, the invention includes a NOVX nucleic acid, wherein one or more nucleotides in the NOVX nucleotide sequence is changed to a different nucleotide provided that no more than 15% of the nucleotides are so changed. In one embodiment, the invention discloses a nucleic acid fragment of the NOVX nucleotide sequence and a nucleic acid fragment wherein one or more nucleotides in the NOVX nucleotide sequence
25 is changed from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed. In another embodiment, the invention includes a nucleic acid molecule wherein the nucleic acid molecule hybridizes under stringent conditions to a NOVX nucleotide sequence or a complement of the NOVX nucleotide sequence. In one embodiment, the invention
30 includes a nucleic acid molecule, wherein the sequence is changed such that no more than 15% of the nucleotides in the coding sequence differ from the NOVX nucleotide sequence or a fragment thereof.

In a further aspect, the invention includes a method for determining the presence or amount of the NOVX nucleic acid in a sample. The method involves the steps of: providing the sample; introducing the sample to a probe that binds to the nucleic acid molecule; and determining the presence or amount of the probe bound to the NOVX
5 nucleic acid molecule, thereby determining the presence or amount of the NOVX nucleic acid molecule in the sample. In one embodiment, the presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type.

In another aspect, the invention discloses a method for determining the presence of or predisposition to a disease associated with altered levels of the NOVX nucleic acid
10 molecule of in a first mammalian subject. The method involves the steps of: measuring the amount of NOVX nucleic acid in a sample from the first mammalian subject; and comparing the amount of the nucleic acid in the sample of step (a) to the amount of NOVX nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease; wherein an alteration in the level of the
15 nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those
20 described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In the case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be
25 limiting.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides novel nucleotides and polypeptides encoded
30 thereby. Included in the invention are the novel nucleic acid sequences, their encoded polypeptides, antibodies, and other related compounds. The sequences are collectively

referred to herein as "NOVX nucleic acids" or "NOVX polynucleotides" and the corresponding encoded polypeptides are referred to as "NOVX polypeptides" or "NOVX proteins." Unless indicated otherwise, "NOVX" is meant to refer to any of the novel sequences disclosed herein. Table 1 provides a summary of the NOVX nucleic acids and their encoded polypeptides.

TABLE 1. Sequences and Corresponding SEQ ID Numbers

NOVX Assignment	Internal Identification	SEQ ID NO (nucleic acid)	SEQ ID NO (polypeptide)	Homology
1a	CG100126-01	1	2	KERATIN ASSOCIATED PROTEIN 4.9
2a	CG100146-01	3	4	UDP-Glucuronosyltransferase 2B15 Precursor like homo sapiens
3a	CG100179-01	5	6	cyclophilin A like homo sapiens
4a	CG100212-01	7	8	Zinc-containing alcohol dehydrogenase like homo sapiens
4b	CG100212-02	9	10	Zinc-containing alcohol dehydrogenase like homo sapiens
5a	CG100222-01	11	12	NADP-Dependent Leukotriene B4 12-Hydroxydehydrogenase like homo sapiens
6a	CG100266-01	13	14	cyclophilin A like homo sapiens
6b	CG100266-02	15	16	cyclophilin A like homo sapiens
7a	CG100427-01	17	18	cyclophilin A like homo sapiens
8a	CG100456-01	19	20	CoA Transferase like homo sapiens
9a	CG100466-01	21	22	Novel Adenine Nucleotide Translocator 2 like homo sapiens
10a	CG100609-01	23	24	GLUTATHIONE S-TRANSFERASE THETA 1 like homo sapiens
11a	CG100631-01	25	26	Clathrin Light Chain A like homo sapiens
11b	CG100631-02	27	28	Clathrin Light Chain A like homo sapiens
11c	CG100631-04	29	30	Clathrin Light Chain A like homo sapiens
12a	CG100710-01	31	32	AAA (ATPase Associated with various Activities) like homo sapiens
13a	CG100730-01	33	34	exoribonuclease like homo sapiens
14a	CG100819-01	35	36	POLYNUCLEOTIDE PHOSPHORYLASE like homo sapiens
15a	CG100872-01	37	38	PROTEIN-ARGININE DEIMINASE TYPE V like

				homo sapiens
16a	CG100980-01	39	40	PROTEIN-ARGININE DEIMINASE TYPE III like homo sapiens
17a	CG172805-01	41	42	Peroxisomal Enoyl-CoA Hydratase like homo sapiens
18a	CG56763-01	43	44	GPCR like homo sapiens
19a	CG56777-01	45	46	Prostaglandin-F synthase 1 like homo sapiens
19b	CG56777-02	47	48	Prostaglandin-F synthase 1 like homo sapiens
20a	CG56941-01	49	50	Ribonuclease H type II like homo sapiens
21a	CG57109-01	51	52	Doublecortin/CAM kinase like homo sapiens
21b	CG57109-02	53	54	Doublecortin/CAM kinase like homo sapiens (also been filed as Cura 691A)
21c	CG57109-03	55	56	Doublecortin/CAM kinase like homo sapiens
21d	CG57109-04	57	58	Doublecortin/CAM kinase like homo sapiens
21e	CG57109-05	59	60	Doublecortin/CAM kinase like homo sapiens
21f	CG57109-06	61	62	Double Cortin
22a	CG57366-01	63	64	KIAA1223 like homo sapiens
23a	CG57368-01	65	66	Adenylate cyclase type IV like homo sapiens
23b	CG57368-02	67	68	
24a	CG59955-01	69	70	GPCR like homo sapiens
25a	CG89211-01	71	72	GPCR like homo sapiens
26a	CG90530-02	73	74	UBIQUITIN-CONJUGATING ENZYME ISOLOG like homo sapiens
27a	CG93076-01	75	76	GPCR like homo sapiens
28a	CG94235-01	77	78	Thymidylate Kinase like homo sapiens
28b	CG94235-02	79	80	Thymidylate Kinase like homo sapiens
29a	CG94692-01	81	82	CARNITINE/ACYLCARNITIN E TRANSLOCASE like homo sapiens
29b	CG94692-02	83	84	CARNITINE/ACYLCARNITIN E TRANSLOCASE like homo sapiens
30a	CG94724-01	85	86	CARNITINE/ACYLCARNITIN E TRANSLOCASE like homo sapiens
31a	CG94871-01	87	88	Josephin MJD1 like homo sapiens
31b	CG94871-02	89	90	Josephin MJD1 like homo sapiens
31c	CG94871-04	91	92	Josephin MJD1 like homo sapiens
31d	CG94871-05	93	94	Josephin MJD1 like homo sapiens
32a	CG94946-01	95	96	Aggrin Precursor

32b	CG94946-02	97	98	Agrin Precursor
32c	CG94946-03	99	100	Agrin Precursor
32d	CG94946-04	101	102	Agrin Precursor
32e	CG94946-05	103	104	Agrin Precursor
32f	CG94946-06	105	106	Agrin Precursor
32g	CG94946-07	107	108	Agrin Precursor
33a	CG95165-01	109	110	ADENYLATE CYCLASE, TYPE II like homo sapiens
34a	CG95175-01	111	112	Ephrine type A Receptor like homo sapiens
35a	CG95693-01	113	114	DPY-19 PROTEIN like homo sapiens
36a	CG95814-01	115	116	RHO-GTPASE-ACTIVATING PROTEIN 4 like homo sapiens
37a	CG95824-01	117	118	Rho GAP like homo sapiens
38a	CG96198-01	119	120	Gonadotropin-releasing hormone receptor
38b	CG96198-02	121	122	Gonadotropin-releasing hormone receptor
39a	CG96231-01	123	124	OTU-like cysteine protease like homo sapiens
39b	CG96231-02	125	126	OTU-like cysteine protease like homo sapiens
40a	CG96260-01	127	128	phospholipase A2 like homo sapiens
41a	CG96364-01	129	130	ADP/ATP TRANSLOCASE 2 like homo sapiens
41b	CG96364-03	131	132	ADP/ATP TRANSLOCASE 2 like homo sapiens
42a	CG96422-01	133	134	ADP/ATP translocase 3 like homo sapiens
43a	CG96442-01	135	136	Acyl CoA-domain protein like homo sapiens
44a	CG96501-01	137	138	BRAIN MITOCHONDRIAL CARRIER PROTEIN-1 like homo sapiens
45a	CG96557-01	139	140	L-ALLO-THREONINE ALDOLASE
46a	CG96581-01	141	142	RP42
46b	CG96581-02	143	144	RP42
47a	CG96624-01	145	146	Putative seven pass transmembrane protein
47b	CG96624-02	147	148	Putative seven pass transmembrane protein
48a	CG96747-01	149	150	VOLTAGE-DEPENDENT CALCIUM CHANNEL GAMMA-3 SUBUNIT
49a	CG96789-01	151	152	GLYCINE CLEAVAGE SYSTEM PROTEIN H
50a	CG97253-01	153	154	Galectin like homo sapiens
51a	CG97400-01	155	156	Myotubularin-related protein 6 like homo sapiens
52a	CG97462-01	157	158	Prohibitin like homo sapiens
53a	CG97472-01	159	160	Glucose Transporter like homo sapiens
54a	CG97528-01	161	162	Guanylate Binding Protein like homo sapiens

55a	CG97629-01	163	164	CELL DIVISION PROTEIN KINASE 7 like homo sapiens
56a	CG97648-01	165	166	G PROTEIN-COUPLED RECEPTOR KINASE GRK7 like homo sapiens
57a	CG97658-01	167	168	PROTEIN-TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 5 like homo sapiens
58a	CG97842-01	169	170	protein kinase-formin like homo sapiens
59a	CG98021-01	171	172	Synaptotagmin III like homo sapiens
60a	CG98030-01	173	174	Protein-tyrosine phosphatase like homo sapiens
60b	CG98030-02	175	176	
61a	CG98061-01	177	178	Histidine acid phosphatase domain like homo sapiens
61b	CG98061-02	179	180	Novel Protein containing Histidine acid phosphatase domain-like Proteins and Nucleic Acids Encoding Same
62a	CG98071-01	181	182	Histidine acid phosphatase domain like homo sapiens
63a	CG98131-01	183	184	MDJ6 like homo sapiens
64a	CG98164-01	185	186	LRR and Kinase domain like homo sapiens
64b	CG98164-02	187	188	LRR and Kinase domain like homo sapiens
65a	CG99588-01	189	190	transmembrane protein like homo sapiens
66a	CG99618-01	191	192	PROTEIN-TYROSINE PHOSPHATASE 2C
67a	CG99832-01	193	194	gene containing NUDIX hydrolase domain like homo sapiens
68a	CG99842-01	195	196	Tensin
69a	CG99944-01	197	198	SUGAR ABC TRANSPORTER
70a	CG99963-01	199	200	cyclophilin 18 like homo sapiens
70b	CG99963-02	201	202	cyclophilin 18 like homo sapiens

Table 1 indicates homology of NOVX nucleic acids to known protein families. Thus, the nucleic acids and polypeptides, antibodies and related compounds according to the invention corresponding to a NOVX as identified in column 1 of Table 1 will be useful in therapeutic and diagnostic applications implicated in, for example, pathologies and disorders associated with the known protein families identified in column 5 of Table 1.

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence of domains and sequence relatedness to previously described proteins.

Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

Consistent with other known members of the family of proteins, identified in column 5 of Table 1, the NOVX polypeptides of the present invention show homology to, and contain domains that are characteristic of, other members of such protein families. Details of the sequence relatedness and domain analysis for each NOVX are presented in Example A.

The NOVX nucleic acids and polypeptides can also be used to screen for molecules, which inhibit or enhance NOVX activity or function. Specifically, the nucleic acids and polypeptides according to the invention may be used as targets for the identification of small molecules that modulate or inhibit diseases associated with the protein families listed in Table 1.

The NOVX nucleic acids and polypeptides are also useful for detecting specific cell types. Details of the expression analysis for each NOVX are presented in Example C. Accordingly, the NOVX nucleic acids, polypeptides, antibodies and related compounds according to the invention will have diagnostic and therapeutic applications in the detection of a variety of diseases with differential expression in normal vs. diseased tissues, e.g. a variety of cancers.

Additional utilities for NOVX nucleic acids and polypeptides according to the invention are disclosed herein.

The present invention is based on the identification of biological macromolecules differentially modulated in a pathologic state, disease, or an abnormal condition or state. Among the pathologies or diseases of present interest include metabolic diseases including those related to endocrinologic disorders, cancers, various tumors and neoplasias, inflammatory disorders, central nervous system disorders, and similar abnormal conditions or states. In very significant embodiments of the present invention, the biological macromolecules implicated in the pathologies and conditions are proteins and polypeptides, and in such cases the present invention is related as well to the nucleic acids that encode them. Methods that may be employed to identify relevant biological macromolecules include any procedures that detect differential expression of nucleic acids encoding proteins and polypeptides associated with the disorder, as well as procedures that detect the respective proteins and polypeptides themselves. Significant methods that have

been employed by the present inventors, include GeneCalling ® technology and SeqCalling TM technology, disclosed respectively, in U. S. Patent No. 5,871,697, and in U. S. Ser. No. 09/417,386, filed Oct. 13, 1999, each of which is incorporated herein by reference in its entirety. GeneCalling ® is also described in Shimkets, et al., “Gene
5 expression analysis by transcript profiling coupled to a gene database query” Nature Biotechnology 17:198-803 (1999).

The invention provides polypeptides and nucleotides encoded thereby that have been identified as having novel associations with a disease or pathology, or an abnormal state or condition, in a mammal. The present invention further identifies a set of proteins
10 and polypeptides, including naturally occurring polypeptides, precursor forms or proproteins, or mature forms of the polypeptides or proteins, which are implicated as targets for therapeutic agents in the treatment of various diseases, pathologies, abnormal states and conditions. A target may be employed in any of a variety of screening methodologies in order to identify candidate therapeutic agents which interact with the
15 target and in so doing exert a desired or favorable effect. The candidate therapeutic agent is identified by screening a large collection of substances or compounds in an important embodiment of the invention. Such a collection may comprise a combinatorial library of substances or compounds in which, in at least one subset of substances or compounds, the individual members are related to each other by simple structural variations based on a
20 particular canonical or basic chemical structure. The variations may include, by way of nonlimiting example, changes in length or identity of a basic framework of bonded atoms; changes in number, composition and disposition of ringed structures, bridge structures, alicyclic rings, and aromatic rings; and changes in pendent or substituents atoms or groups that are bonded at particular positions to the basic framework of bonded atoms or to the
25 ringed structures, the bridge structures, the alicyclic structures, or the aromatic structures.

A polypeptide or protein described herein, and that serves as a target in the screening procedure, includes the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, e.g., the full-length gene product, encoded by the corresponding gene. The
30 naturally occurring polypeptide also includes the polypeptide, precursor or proprotein encoded by an open reading frame described herein. A “mature” form of a polypeptide or protein arises as a result of one or more naturally occurring processing steps as they may

occur within the cell, including a host cell. The processing steps occur as the gene product arises, e.g., via cleavage of the amino-terminal methionine residue encoded by the initiation codon of an open reading frame, or the proteolytic cleavage of a signal peptide or leader sequence. Thus, a mature form arising from a precursor polypeptide or protein
5 that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining. Alternatively, a mature form arising from a precursor polypeptide or protein having residues 1 to N, in which an amino-terminal signal sequence from residue 1 to residue M is cleaved, includes the residues from residue M+1 to residue N remaining. A "mature" form of a polypeptide or protein may also arise from non-
10 proteolytic post-translational modification. Such non-proteolytic processes include, e.g., glycosylation, myristylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or the combination of any of them.

As used herein, "identical" residues correspond to those residues in a comparison
15 between two sequences where the equivalent nucleotide base or amino acid residue in an alignment of two sequences is the same residue. Residues are alternatively described as "similar" or "positive" when the comparisons between two sequences in an alignment show that residues in an equivalent position in a comparison are either the same amino acid or a conserved amino acid as defined below.

20 As used herein, a "chemical composition" relates to a composition including at least one compound that is either synthesized or extracted from a natural source. A chemical compound may be the product of a defined synthetic procedure. Such a synthesized compound is understood herein to have defined properties in terms of molecular formula, molecular structure relating the association of bonded atoms to each
25 other, physical properties such as chromatographic or spectroscopic characterizations, and the like. A compound extracted from a natural source is advantageously analyzed by chemical and physical methods in order to provide a representation of its defined properties, including its molecular formula, molecular structure relating the association of bonded atoms to each other, physical properties such as chromatographic or spectroscopic
30 characterizations, and the like.

As used herein, a "candidate therapeutic agent" is a chemical compound that includes at least one substance shown to bind to a target biopolymer. In important

embodiments of the invention, the target biopolymer is a protein or polypeptide, a nucleic acid, a polysaccharide or proteoglycan, or a lipid such as a complex lipid. The method of identifying compounds that bind to the target effectively eliminates compounds with little or no binding affinity, thereby increasing the potential that the identified chemical
5 compound may have beneficial therapeutic applications. In cases where the "candidate therapeutic agent" is a mixture of more than one chemical compound, subsequent screening procedures may be carried out to identify the particular substance in the mixture that is the binding compound, and that is to be identified as a candidate therapeutic agent.

As used herein, a "pharmaceutical agent" is provided by screening a candidate
10 therapeutic agent using models for a disease state or pathology in order to identify a candidate exerting a desired or beneficial therapeutic effect with relation to the disease or pathology. Such a candidate that successfully provides such an effect is termed a pharmaceutical agent herein. Nonlimiting examples of model systems that may be used in such screens include particular cell lines, cultured cells, tissue preparations, whole tissues,
15 organ preparations, intact organs, and nonhuman mammals. Screens employing at least one system, and preferably more than one system, may be employed in order to identify a pharmaceutical agent. Any pharmaceutical agent so identified may be pursued in further investigation using human subjects.

NOVX Nucleic Acids and Polypeptides

20 NOVX clones

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence of domains and sequence relatedness to previously described proteins.
25 Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

The NOVX genes and their corresponding encoded proteins are useful for preventing, treating or ameliorating medical conditions, *e.g.*, by protein or gene therapy. Pathological conditions can be diagnosed by determining the amount of the new protein in
30 a sample or by determining the presence of mutations in the new genes. Specific uses are described for each of the NOVX genes, based on the tissues in which they are most highly

expressed. Uses include developing products for the diagnosis or treatment of a variety of diseases and disorders.

The NOVX nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications and as a research tool. These include serving as a
5 specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene
10 delivery/gene ablation), and (v) a composition promoting tissue regeneration *in vitro* and *in vivo* (vi) biological defense weapon.

In one specific embodiment, the invention includes an isolated polypeptide comprising an amino acid sequence selected from the group consisting of: (a) a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n,
15 wherein n is an integer between 1 and 101; (b) a variant of a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101, wherein any amino acid in the mature form is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence of the mature form are so changed; (c) an amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101; (d) a
20 variant of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101 wherein any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed; and (e) a fragment of any of (a)
25 through (d).

In another specific embodiment, the invention includes an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of: (a) a mature form of the amino acid sequence given SEQ ID NO: 2n, wherein n is an integer between 1 and 101; (b) a variant
30 of a mature form of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101 wherein any amino acid in the mature form of the chosen sequence is changed to a different amino acid, provided that no

more than 15% of the amino acid residues in the sequence of the mature form are so changed; (c) the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101; (d) a variant of the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1
5 and 101, in which any amino acid specified in the chosen sequence is changed to a different amino acid, provided that no more than 15% of the amino acid residues in the sequence are so changed; (e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO: 2n, wherein n is an integer between 1 and 101 or any variant of said
10 polypeptide wherein any amino acid of the chosen sequence is changed to a different amino acid, provided that no more than 10% of the amino acid residues in the sequence are so changed; and (f) the complement of any of said nucleic acid molecules.

In yet another specific embodiment, the invention includes an isolated nucleic acid molecule, wherein said nucleic acid molecule comprises a nucleotide sequence selected
15 from the group consisting of: (a) the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101; (b) a nucleotide sequence wherein one or more nucleotides in the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 is changed from that selected from the group consisting of the chosen sequence to a
20 different nucleotide provided that no more than 15% of the nucleotides are so changed; (c) a nucleic acid fragment of the sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101; and (d) a nucleic acid fragment wherein one or more nucleotides in the nucleotide sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 is changed
25 from that selected from the group consisting of the chosen sequence to a different nucleotide provided that no more than 15% of the nucleotides are so changed.

One aspect of the invention pertains to isolated nucleic acid molecules that encode NOVX polypeptides or biologically active portions thereof. Also included in the invention are nucleic acid fragments sufficient for use as hybridization probes to identify
30 NOVX-encoding nucleic acids (*e.g.*, NOVX mRNAs) and fragments for use as PCR primers for the amplification and/or mutation of NOVX nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*,

cDNA or genomic DNA), RNA molecules (*e.g.*, mRNA), analogs of the DNA or RNA generated using nucleotide analogs, and derivatives, fragments and homologs thereof. The nucleic acid molecule may be single-stranded or double-stranded, but preferably is comprised double-stranded DNA.

5 An NOVX nucleic acid can encode a mature NOVX polypeptide. As used herein, a "mature" form of a polypeptide or protein disclosed in the present invention is the product of a naturally occurring polypeptide or precursor form or proprotein. The naturally occurring polypeptide, precursor or proprotein includes, by way of nonlimiting example, the full-length gene product, encoded by the corresponding gene. Alternatively,
10 it may be defined as the polypeptide, precursor or proprotein encoded by an ORF described herein. The product "mature" form arises, again by way of nonlimiting example, as a result of one or more naturally occurring processing steps as they may take place within the cell, or host cell, in which the gene product arises. Examples of such processing steps leading to a "mature" form of a polypeptide or protein include the
15 cleavage of the N-terminal methionine residue encoded by the initiation codon of an ORF, or the proteolytic cleavage of a signal peptide or leader sequence. Thus a mature form arising from a precursor polypeptide or protein that has residues 1 to N, where residue 1 is the N-terminal methionine, would have residues 2 through N remaining after removal of the N-terminal methionine. Alternatively, a mature form arising from a precursor
20 polypeptide or protein having residues 1 to N, in which an N-terminal signal sequence from residue 1 to residue M is cleaved, would have the residues from residue M+1 to residue N remaining. Further as used herein, a "mature" form of a polypeptide or protein may arise from a step of post-translational modification other than a proteolytic cleavage event. Such additional processes include, by way of non-limiting example, glycosylation,
25 myristoylation or phosphorylation. In general, a mature polypeptide or protein may result from the operation of only one of these processes, or a combination of any of them.

 The term "probes", as utilized herein, refers to nucleic acid sequences of variable length, preferably between at least about 10 nucleotides (nt), 100 nt, or as many as approximately, *e.g.*, 6,000 nt, depending upon the specific use. Probes are used in the
30 detection of identical, similar, or complementary nucleic acid sequences. Longer length probes are generally obtained from a natural or recombinant source, are highly specific, and much slower to hybridize than shorter-length oligomer probes. Probes may be single-

or double-stranded and designed to have specificity in PCR, membrane-based hybridization technologies, or ELISA-like technologies.

The term "isolated" nucleic acid molecule, as utilized herein, is one, which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5'- and 3'-termini of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated NOVX nucleic acid molecules can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell/tissue from which the nucleic acid is derived (*e.g.*, brain, heart, liver, spleen, etc.). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material or culture medium when produced by recombinant techniques, or of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or a complement of this aforementioned nucleotide sequence, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 as a hybridization probe, NOVX molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, *et al.*, (eds.), MOLECULAR CLONING: A LABORATORY MANUAL 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993.)

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to NOVX nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

As used herein, the term "oligonucleotide" refers to a series of linked nucleotide residues, which oligonucleotide has a sufficient number of nucleotide bases to be used in a

PCR reaction. A short oligonucleotide sequence may be based on, or designed from, a genomic or cDNA sequence and is used to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA or RNA in a particular cell or tissue.

Oligonucleotides comprise portions of a nucleic acid sequence having about 10 nt, 50 nt, or 100 nt in length, preferably about 15 nt to 30 nt in length. In one embodiment of the invention, an oligonucleotide comprising a nucleic acid molecule less than 100 nt in length would further comprise at least 6 contiguous nucleotides SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or a complement thereof. Oligonucleotides may be chemically synthesized and may also be used as probes.

10 In another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule that is a complement of the nucleotide from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or a portion of this nucleotide sequence (e.g., a fragment that can be used as a probe or primer or a fragment encoding a biologically-active portion of an NOVX polypeptide). A nucleic acid molecule that is complementary to the nucleotide sequence from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 is one that is sufficiently complementary to the nucleotide sequence from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 that it can hydrogen bond with little or no mismatches to the nucleotide sequence from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, thereby forming a stable duplex.

As used herein, the term "complementary" refers to Watson-Crick or Hoogsteen base pairing between nucleotides units of a nucleic acid molecule, and the term "binding" means the physical or chemical interaction between two polypeptides or compounds or associated polypeptides or compounds or combinations thereof. Binding includes ionic, non-ionic, van der Waals, hydrophobic interactions, and the like. A physical interaction can be either direct or indirect. Indirect interactions may be through or due to the effects of another polypeptide or compound. Direct binding refers to interactions that do not take place through, or due to, the effect of another polypeptide or compound, but instead are without other substantial chemical intermediates.

30 Fragments provided herein are defined as sequences of at least 6 (contiguous) nucleic acids or at least 4 (contiguous) amino acids, a length sufficient to allow for specific hybridization in the case of nucleic acids or for specific recognition of an epitope

in the case of amino acids, respectively, and are at most some portion less than a full length sequence. Fragments may be derived from any contiguous portion of a nucleic acid or amino acid sequence of choice. Derivatives are nucleic acid sequences or amino acid sequences formed from the native compounds either directly or by modification or partial substitution. Analogs are nucleic acid sequences or amino acid sequences that have a structure similar to, but not identical to, the native compound but differs from it in respect to certain components or side chains. Analogs may be synthetic or from a different evolutionary origin and may have a similar or opposite metabolic activity compared to wild type. Homologs are nucleic acid sequences or amino acid sequences of a particular gene that are derived from different species.

A full-length NOVX clone is identified as containing an ATG translation start codon and an in-frame stop codon. Any disclosed NOVX nucleotide sequence lacking an ATG start codon therefore encodes a truncated C-terminal fragment of the respective NOVX polypeptide, and requires that the corresponding full-length cDNA extend in the 5' direction of the disclosed sequence. Any disclosed NOVX nucleotide sequence lacking an in-frame stop codon similarly encodes a truncated N-terminal fragment of the respective NOVX polypeptide, and requires that the corresponding full-length cDNA extend in the 3' direction of the disclosed sequence.

Derivatives and analogs may be full length or other than full length, if the derivative or analog contains a modified nucleic acid or amino acid, as described below. Derivatives or analogs of the nucleic acids or proteins of the invention include, but are not limited to, molecules comprising regions that are substantially homologous to the nucleic acids or proteins of the invention, in various embodiments, by at least about 70%, 80%, or 95% identity (with a preferred identity of 80-95%) over a nucleic acid or amino acid sequence of identical size or when compared to an aligned sequence in which the alignment is done by a computer homology program known in the art, or whose encoding nucleic acid is capable of hybridizing to the complement of a sequence encoding the aforementioned proteins under stringent, moderately stringent, or low stringent conditions. See *e.g.* Ausubel, *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, NY, 1993, and below.

A "homologous nucleic acid sequence" or "homologous amino acid sequence," or variations thereof, refer to sequences characterized by a homology at the nucleotide level

or amino acid level as discussed above. Homologous nucleotide sequences encode those sequences coding for isoforms of NOVX polypeptides. Isoforms can be expressed in different tissues of the same organism as a result of, for example, alternative splicing of RNA. Alternatively, isoforms can be encoded by different genes. In the invention,

5 homologous nucleotide sequences include nucleotide sequences encoding for an NOVX polypeptide of species other than humans, including, but not limited to: vertebrates, and thus can include, *e.g.*, frog, mouse, rat, rabbit, dog, cat, cow, horse, and other organisms. Homologous nucleotide sequences also include, but are not limited to, naturally occurring allelic variations and mutations of the nucleotide sequences set forth herein. A

10 homologous nucleotide sequence does not, however, include the exact nucleotide sequence encoding human NOVX protein. Homologous nucleic acid sequences include those nucleic acid sequences that encode conservative amino acid substitutions (see below) in SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, as well as a polypeptide possessing NOVX biological activity. Various biological activities of the

15 NOVX proteins are described below.

An NOVX polypeptide is encoded by the open reading frame ("ORF") of an NOVX nucleic acid. An ORF corresponds to a nucleotide sequence that could potentially be translated into a polypeptide. A stretch of nucleic acids comprising an ORF is uninterrupted by a stop codon. An ORF that represents the coding sequence for a full

20 protein begins with an ATG "start" codon and terminates with one of the three "stop" codons, namely, TAA, TAG, or TGA. For the purposes of this invention, an ORF may be any part of a coding sequence, with or without a start codon, a stop codon, or both. For an ORF to be considered as a good candidate for coding for a *bona fide* cellular protein, a minimum size requirement is often set, *e.g.*, a stretch of DNA that would encode a protein

25 of 50 amino acids or more.

The nucleotide sequences determined from the cloning of the human NOVX genes allows for the generation of probes and primers designed for use in identifying and/or cloning NOVX homologues in other cell types, *e.g.* from other tissues, as well as NOVX homologues from other vertebrates. The probe/primer typically comprises substantially

30 purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, 25, 50, 100, 150, 200, 250, 300, 350 or 400 consecutive sense strand nucleotide sequence SEQ ID NO: 2n-

1, wherein n is an integer between 1 and 101; or an anti-sense strand nucleotide sequence of SEQ ID NO: $2n-1$, wherein n is an integer between 1 and 101.

Probes based on the human NOVX nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In various
5 embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissues which mis-express an NOVX protein, such as by measuring a level of an NOVX-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting NOVX mRNA levels or
10 determining whether a genomic NOVX gene has been mutated or deleted.

"A polypeptide having a biologically-active portion of an NOVX polypeptide" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. A nucleic acid fragment
15 encoding a "biologically-active portion of NOVX" can be prepared by isolating a portion SEQ ID NO: $2n-1$, wherein n is an integer between 1 and 101, that encodes a polypeptide having an NOVX biological activity (the biological activities of the NOVX proteins are described below), expressing the encoded portion of NOVX protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of NOVX.

20 NOVX Nucleic Acid and Polypeptide Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequences shown in SEQ ID NO: $2n-1$, wherein n is an integer between 1 and 101 due to degeneracy of the genetic code and thus encode the same NOVX proteins as that encoded by the nucleotide sequences shown in SEQ ID NO: $2n-1$, wherein n is an
25 integer between 1 and 101. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in SEQ ID NO: $2n$, wherein n is an integer between 1 and 101.

In addition to the human NOVX nucleotide sequences shown in SEQ ID NO: $2n-1$, wherein n is an integer between 1 and 101, it will be appreciated by those skilled in the art
30 that DNA sequence polymorphisms that lead to changes in the amino acid sequences of the NOVX polypeptides may exist within a population (*e.g.*, the human population). Such genetic polymorphism in the NOVX genes may exist among individuals within a

population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame (ORF) encoding an NOVX protein, preferably a vertebrate NOVX protein. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of the NOVX genes. Any and all such nucleotide variations and resulting amino acid polymorphisms in the NOVX polypeptides, which are the result of natural allelic variation and that do not alter the functional activity of the NOVX polypeptides, are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding NOVX proteins from other species, and thus that have a nucleotide sequence that differs from the human SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of the NOVX cDNAs of the invention can be isolated based on their homology to the human NOVX nucleic acids disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 6 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101. In another embodiment, the nucleic acid is at least 10, 25, 50, 100, 250, 500, 750, 1000, 1500, or 2000 or more nucleotides in length. In yet another embodiment, an isolated nucleic acid molecule of the invention hybridizes to the coding region. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other.

Homologs (*i.e.*, nucleic acids encoding NOVX proteins derived from species other than human) or other related sequences (*e.g.*, paralogs) can be obtained by low, moderate or high stringency hybridization with all or a portion of the particular human sequence as a probe using methods well known in the art for nucleic acid hybridization and cloning.

As used herein, the phrase "stringent hybridization conditions" refers to conditions under which a probe, primer or oligonucleotide will hybridize to its target sequence, but to

no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures than shorter sequences. Generally, stringent conditions are selected to be about 5 °C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength, pH and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Since the target sequences are generally present at excess, at T_m , 50% of the probes are occupied at equilibrium. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes, primers or oligonucleotides (*e.g.*, 10 nt to 50 nt) and at least about 60°C for longer probes, primers and oligonucleotides. Stringent conditions may also be achieved with the addition of destabilizing agents, such as formamide.

Stringent conditions are known to those skilled in the art and can be found in Ausubel, *et al.*, (eds.), CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, N.Y. (1989). 6.3.1-6.3.6. Preferably, the conditions are such that sequences at least about 65%, 70%, 75%, 85%, 90%, 95%, 98%, or 99% homologous to each other typically remain hybridized to each other. A non-limiting example of stringent hybridization conditions are hybridization in a high salt buffer comprising 6X SSC, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.02% BSA, and 500 mg/ml denatured salmon sperm DNA at 65°C, followed by one or more washes in 0.2X SSC, 0.01% BSA at 50°C. An isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequences SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In a second embodiment, a nucleic acid sequence that is hybridizable to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or fragments, analogs or derivatives thereof, under conditions of moderate stringency is provided. A non-limiting example of moderate stringency hybridization conditions are hybridization in 6X SSC, 5X Denhardt's solution, 0.5% SDS

and 100 mg/ml denatured salmon sperm DNA at 55°C, followed by one or more washes in 1X SSC, 0.1% SDS at 37°C. Other conditions of moderate stringency that may be used are well-known within the art. *See, e.g.,* Ausubel, et al. (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990; GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY.

In a third embodiment, a nucleic acid that is hybridizable to the nucleic acid molecule comprising the nucleotide sequences SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or fragments, analogs or derivatives thereof, under conditions of low stringency, is provided. A non-limiting example of low stringency hybridization conditions are hybridization in 35% formamide, 5X SSC, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, 0.02% PVP, 0.02% Ficoll, 0.2% BSA, 100 mg/ml denatured salmon sperm DNA, 10% (wt/vol) dextran sulfate at 40°C, followed by one or more washes in 2X SSC, 25 mM Tris-HCl (pH 7.4), 5 mM EDTA, and 0.1% SDS at 50°C. Other conditions of low stringency that may be used are well known in the art (*e.g.,* as employed for cross-species hybridizations). *See, e.g.,* Ausubel, et al. (eds.), 1993, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, NY, and Kriegler, 1990, GENE TRANSFER AND EXPRESSION, A LABORATORY MANUAL, Stockton Press, NY; Shilo and Weinberg, 1981. *Proc Natl Acad Sci USA* 78: 6789-6792.

Conservative Mutations

In addition to naturally-occurring allelic variants of NOVX sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, thereby leading to changes in the amino acid sequences of the encoded NOVX proteins, without altering the functional ability of said NOVX proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence SEQ ID NO: 2n, wherein n is an integer between 1 and 101. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequences of the NOVX proteins without altering their biological activity, whereas an "essential" amino acid residue is required for such biological activity. For example, amino acid residues that are conserved among the NOVX proteins of the invention are predicted to be particularly non-amenable to

alteration. Amino acids for which conservative substitutions can be made are well-known within the art.

Another aspect of the invention pertains to nucleic acid molecules encoding NOVX proteins that contain changes in amino acid residues that are not essential for activity. Such NOVX proteins differ in amino acid sequence from SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 45% homologous to the amino acid sequences SEQ ID NO: 2n, wherein n is an integer between 1 and 101. Preferably, the protein encoded by the nucleic acid molecule is at least about 60% homologous to SEQ ID NO: 2n, wherein n is an integer between 1 and 101; more preferably at least about 70% homologous SEQ ID NO: 2n, wherein n is an integer between 1 and 101; still more preferably at least about 80% homologous to SEQ ID NO: 2n, wherein n is an integer between 1 and 101; even more preferably at least about 90% homologous to SEQ ID NO: 2n, wherein n is an integer between 1 and 101; and most preferably at least about 95% homologous to SEQ ID NO: 2n, wherein n is an integer between 1 and 101.

An isolated nucleic acid molecule encoding an NOVX protein homologous to the protein of SEQ ID NO: 2n, wherein n is an integer between 1 and 101 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein.

Mutations can be introduced into SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted, non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined within the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine,

tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted non-essential amino acid residue in the NOVX protein is replaced with
5 another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an NOVX coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for NOVX biological activity to identify mutants that retain activity. Following mutagenesis
10 SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, the encoded protein can be expressed by any recombinant technology known in the art and the activity of the protein can be determined.

The relatedness of amino acid families may also be determined based on side chain interactions. Substituted amino acids may be fully conserved "strong" residues or fully conserved "weak" residues. The "strong" group of conserved amino acid residues may be
15 any one of the following groups: STA, NEQK, NHQK, NDEQ, QHRK, MILV, MILF, HY, FYW, wherein the single letter amino acid codes are grouped by those amino acids that may be substituted for each other. Likewise, the "weak" group of conserved residues may be any one of the following: CSA, ATV, SAG, STNK, STPA, SGND, SNDEQK, NDEQHK, NEQHRK, HFY, wherein the letters within each group represent the single
20 letter amino acid code.

In one embodiment, a mutant NOVX protein can be assayed for (i) the ability to form protein:protein interactions with other NOVX proteins, other cell-surface proteins, or biologically-active portions thereof, (ii) complex formation between a mutant NOVX protein and an NOVX ligand; or (iii) the ability of a mutant NOVX protein to bind to an
25 intracellular target protein or biologically-active portion thereof; (*e.g.* avidin proteins).

In yet another embodiment, a mutant NOVX protein can be assayed for the ability to regulate a specific biological function (*e.g.*, regulation of insulin release).

Antisense Nucleic Acids

Another aspect of the invention pertains to isolated antisense nucleic acid
30 molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid

comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein (*e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence). In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25,
5 50, 100, 250 or 500 nucleotides or an entire NOVX coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of an NOVX protein of SEQ ID NO: 2n, wherein n is an integer between 1 and 101, or antisense nucleic acids complementary to an NOVX nucleic acid sequence of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, are additionally provided.

10 In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an NOVX protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a
15 nucleotide sequence encoding the NOVX protein. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding the NOVX protein disclosed herein, antisense nucleic acids of the invention can be designed according to the rules of Watson
20 and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of NOVX mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of NOVX mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of NOVX mRNA. An antisense
25 oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally-occurring nucleotides or variously modified nucleotides
30 designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids (*e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used).

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an NOVX protein to thereby inhibit expression of the protein (*e.g.*, by inhibiting transcription and/or translation). The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface (*e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens). The antisense nucleic acid molecules can also be delivered to cells using the

vectors described herein. To achieve sufficient nucleic acid molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is
5 an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other. *See, e.g., Gaultier, et al., 1987. Nucl. Acids Res. 15: 6625-6641.* The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (*See, e.g., Inoue, et al. 1987. Nucl. Acids Res. 15: 6131-6148*)
10 or a chimeric RNA-DNA analogue (*See, e.g., Inoue, et al., 1987. FEBS Lett. 215: 327-330.*

Ribozymes and PNA Moieties

Nucleic acid modifications include, by way of non-limiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized.
15 These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be used, for example, as antisense binding nucleic acids in therapeutic applications in a subject.

In one embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of
20 cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g., hammerhead ribozymes as described in Haselhoff and Gerlach 1988. Nature 334: 585-591*) can be used to catalytically cleave NOVX mRNA transcripts to thereby inhibit translation of NOVX mRNA. A ribozyme having specificity for an NOVX-encoding nucleic acid can be designed based upon the
25 nucleotide sequence of an NOVX cDNA disclosed herein (*i.e., SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101*). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an NOVX-encoding mRNA. *See, e.g., U.S. Patent 4,987,071 to Cech, et al. and U.S. Patent 5,116,742 to Cech, et al.*
30 NOVX mRNA can also be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g., Bartel et al., (1993) Science 261:1411-1418.*

Alternatively, NOVX gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the NOVX nucleic acid (e.g., the NOVX promoter and/or enhancers) to form triple helical structures that prevent transcription of the NOVX gene in target cells. See, e.g., Helene, 1991. *Anticancer Drug Des.* 6: 569-84; Helene, *et al.* 1992. *Ann. N.Y. Acad. Sci.* 660: 27-36; Maher, 1992. *Bioassays* 14: 807-15.

In various embodiments, the NOVX nucleic acids can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids. See, e.g., Hyrup, *et al.*, 1996. *Bioorg Med Chem* 4: 5-23. As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics (e.g., DNA mimics) in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup, *et al.*, 1996. *supra*; Perry-O'Keefe, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93: 14670-14675.

PNAs of NOVX can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of NOVX can also be used, for example, in the analysis of single base pair mutations in a gene (e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S₁ nucleases (See, Hyrup, *et al.*, 1996. *supra*); or as probes or primers for DNA sequence and hybridization (See, Hyrup, *et al.*, 1996. *supra*; Perry-O'Keefe, *et al.*, 1996. *supra*).

In another embodiment, PNAs of NOVX can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA-chimeras of NOVX can be generated that may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes (e.g., RNase H and DNA polymerases) to interact with the

DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (*see*, Hyrup, et al., 1996. *supra*). The synthesis of PNA-DNA chimeras can be performed as described in

5 Hyrup, *et al.*, 1996. *supra* and Finn, *et al.*, 1996. *Nucl Acids Res* 24: 3357-3363. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA. *See, e.g.*, Mag, *et al.*, 1989. *Nucl Acid Res* 17: 5973-5988.

10 PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment. *See, e.g.*, Finn, *et al.*, 1996. *supra*. Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. *See, e.g.*, Petersen, *et al.*, 1975. *Bioorg. Med. Chem. Lett.* 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups

15 such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (*see, e.g.*, Letsinger, *et al.*, 1989. *Proc. Natl. Acad. Sci. U.S.A.* 86: 6553-6556; Lemaitre, *et al.*, 1987. *Proc. Natl. Acad. Sci.* 84: 648-652; PCT Publication No. WO88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. WO 89/10134). In addition, oligonucleotides can be modified with hybridization triggered

20 cleavage agents (*see, e.g.*, Krol, *et al.*, 1988. *BioTechniques* 6:958-976) or intercalating agents (*see, e.g.*, Zon, 1988. *Pharm. Res.* 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, *e.g.*, a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, and the like.

25 NO VX Polypeptides

A polypeptide according to the invention includes a polypeptide including the amino acid sequence of NO VX polypeptides whose sequences are provided in SEQ ID NO: 2n, wherein n is an integer between 1 and 101. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residues

30 shown in SEQ ID NO: 2n, wherein n is an integer between 1 and 101 while still encoding a protein that maintains its NO VX activities and physiological functions, or a functional fragment thereof.

In general, an NOVX variant that preserves NOVX-like function includes any variant in which residues at a particular position in the sequence have been substituted by other amino acids, and further include the possibility of inserting an additional residue or residues between two residues of the parent protein as well as the possibility of deleting
5 one or more residues from the parent sequence. Any amino acid substitution, insertion, or deletion is encompassed by the invention. In favorable circumstances, the substitution is a conservative substitution as defined above.

One aspect of the invention pertains to isolated NOVX proteins, and biologically-active portions thereof, or derivatives, fragments, analogs or homologs thereof. Also
10 provided are polypeptide fragments suitable for use as immunogens to raise anti-NOVX antibodies. In one embodiment, native NOVX proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, NOVX proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, an NOVX protein or polypeptide can
15 be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" polypeptide or protein or biologically-active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the NOVX protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language
20 "substantially free of cellular material" includes preparations of NOVX proteins in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly-produced. In one embodiment, the language "substantially free of cellular material" includes preparations of NOVX proteins having less than about 30% (by dry weight) of non-NOVX proteins (also referred to herein as a "contaminating protein"),
25 more preferably less than about 20% of non-NOVX proteins, still more preferably less than about 10% of non-NOVX proteins, and most preferably less than about 5% of non-NOVX proteins. When the NOVX protein or biologically-active portion thereof is recombinantly-produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and
30 most preferably less than about 5% of the volume of the NOVX protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of NOVX proteins in which the protein is separated from chemical

precursors or other chemicals that are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of NOVX proteins having less than about 30% (by dry weight) of chemical precursors or non-NOVX chemicals, more preferably less than about 20%
5 chemical precursors or non-NOVX chemicals, still more preferably less than about 10% chemical precursors or non-NOVX chemicals, and most preferably less than about 5% chemical precursors or non-NOVX chemicals.

Biologically-active portions of NOVX proteins include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequences of
10 the NOVX proteins (*e.g.*, the amino acid sequence shown in SEQ ID NO: 2n, wherein n is an integer between 1 and 101) that include fewer amino acids than the full-length NOVX proteins, and exhibit at least one activity of an NOVX protein. Typically, biologically-active portions comprise a domain or motif with at least one activity of the NOVX protein. A biologically-active portion of an NOVX protein can be a polypeptide which is, for
15 example, 10, 25, 50, 100 or more amino acid residues in length. Moreover, other biologically-active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native NOVX protein.

In an embodiment, the NOVX protein has an amino acid sequence shown SEQ ID
20 NO: 2n, wherein n is an integer between 1 and 101. In other embodiments, the NOVX protein is substantially homologous to SEQ ID NO: 2n, wherein n is an integer between 1 and 101, and retains the functional activity of the protein of SEQ ID NO: 2n, wherein n is an integer between 1 and 101, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail, below. Accordingly, in another
25 embodiment, the NOVX protein is a protein that comprises an amino acid sequence at least about 45% homologous to the amino acid sequence SEQ ID NO: 2n, wherein n is an integer between 1 and 101, and retains the functional activity of the NOVX proteins of SEQ ID NO: 2n, wherein n is an integer between 1 and 101.

Determining Homology Between Two or More Sequences

30 To determine the percent homology of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal

alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the
5 molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity").

The nucleic acid sequence homology may be determined as the degree of identity between two sequences. The homology may be determined using computer programs known in the art, such as GAP software provided in the GCG program package. *See*, Needleman and
10 Wunsch, 1970. *J Mol Biol* 48: 443-453. Using GCG GAP software with the following settings for nucleic acid sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3, the coding region of the analogous nucleic acid sequences referred to above exhibits a degree of identity preferably of at least 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99%, with the CDS (encoding) part of the DNA from the group
15 consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101.

The term "sequence identity" refers to the degree to which two polynucleotide or polypeptide sequences are identical on a residue-by-residue basis over a particular region of comparison. The term "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over that region of comparison, determining the number
20 of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, U, or I, in the case of nucleic acids) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the region of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The term "substantial identity" as used herein denotes a
25 characteristic of a polynucleotide sequence, wherein the polynucleotide comprises a sequence that has at least 80 percent sequence identity, preferably at least 85 percent identity and often 90 to 95 percent sequence identity, more usually at least 99 percent sequence identity as compared to a reference sequence over a comparison region.

Chimeric and Fusion Proteins

30 The invention also provides NOVX chimeric or fusion proteins. As used herein, an NOVX "chimeric protein" or "fusion protein" comprises an NOVX polypeptide operatively-linked to a non-NOVX polypeptide. An "NOVX polypeptide" refers to a

polypeptide having an amino acid sequence corresponding to an NOVX protein SEQ ID NO: 2n, wherein n is an integer between 1 and 101, whereas a "non-NOVX polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein that is not substantially homologous to the NOVX protein, *e.g.*, a protein that is different from the NOVX protein and that is derived from the same or a different organism. Within an NOVX fusion protein the NOVX polypeptide can correspond to all or a portion of an NOVX protein. In one embodiment, an NOVX fusion protein comprises at least one biologically-active portion of an NOVX protein. In another embodiment, an NOVX fusion protein comprises at least two biologically-active portions of an NOVX protein. In yet another embodiment, an NOVX fusion protein comprises at least three biologically-active portions of an NOVX protein. Within the fusion protein, the term "operatively-linked" is intended to indicate that the NOVX polypeptide and the non-NOVX polypeptide are fused in-frame with one another. The non-NOVX polypeptide can be fused to the N-terminus or C-terminus of the NOVX polypeptide.

In one embodiment, the fusion protein is a GST-NOVX fusion protein in which the NOVX sequences are fused to the C-terminus of the GST (glutathione S-transferase) sequences. Such fusion proteins can facilitate the purification of recombinant NOVX polypeptides.

In another embodiment, the fusion protein is an NOVX protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of NOVX can be increased through use of a heterologous signal sequence.

In yet another embodiment, the fusion protein is an NOVX-immunoglobulin fusion protein in which the NOVX sequences are fused to sequences derived from a member of the immunoglobulin protein family. The NOVX-immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between an NOVX ligand and an NOVX protein on the surface of a cell, to thereby suppress NOVX-mediated signal transduction *in vivo*. The NOVX-immunoglobulin fusion proteins can be used to affect the bioavailability of an NOVX cognate ligand. Inhibition of the NOVX ligand/NOVX interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, as well as modulating (*e.g.* promoting or inhibiting) cell survival. Moreover, the

NOVX-immunoglobulin fusion proteins of the invention can be used as immunogens to produce anti-NOVX antibodies in a subject, to purify NOVX ligands, and in screening assays to identify molecules that inhibit the interaction of NOVX with an NOVX ligand.

5 An NOVX chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, *e.g.*, by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic
10 ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.*, Ausubel, *et al.* (eds.)
15 CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An NOVX-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the NOVX protein.

20 NOVX Agonists and Antagonists

The invention also pertains to variants of the NOVX proteins that function as either NOVX agonists (*i.e.*, mimetics) or as NOVX antagonists. Variants of the NOVX protein can be generated by mutagenesis (*e.g.*, discrete point mutation or truncation of the NOVX
25 protein). An agonist of the NOVX protein can retain substantially the same, or a subset of, the biological activities of the naturally occurring form of the NOVX protein. An antagonist of the NOVX protein can inhibit one or more of the activities of the naturally occurring form of the NOVX protein by, for example, competitively binding to a downstream or upstream member of a cellular signaling cascade which includes the
30 NOVX protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein

has fewer side effects in a subject relative to treatment with the naturally occurring form of the NOVX proteins.

Variants of the NOVX proteins that function as either NOVX agonists (*i.e.*, mimetics) or as NOVX antagonists can be identified by screening combinatorial libraries of mutants (*e.g.*, truncation mutants) of the NOVX proteins for NOVX protein agonist or antagonist activity. In one embodiment, a variegated library of NOVX variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of NOVX variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential NOVX sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of NOVX sequences therein. There are a variety of methods which can be used to produce libraries of potential NOVX variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential NOVX sequences. Methods for synthesizing degenerate oligonucleotides are well-known within the art. *See, e.g.*, Narang, 1983. *Tetrahedron* 39: 3; Itakura, *et al.*, 1984. *Annu. Rev. Biochem.* 53: 323; Itakura, *et al.*, 1984. *Science* 198: 1056; Ike, *et al.*, 1983. *Nucl. Acids Res.* 11: 477.

Polypeptide Libraries

In addition, libraries of fragments of the NOVX protein coding sequences can be used to generate a variegated population of NOVX fragments for screening and subsequent selection of variants of an NOVX protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an NOVX coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double-stranded DNA that can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S_1 nuclease, and ligating the resulting fragment library into an expression vector. By this

method, expression libraries can be derived which encodes N-terminal and internal fragments of various sizes of the NOVX proteins.

Various techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA
5 libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of NOVX proteins. The most widely used techniques, which are amenable to high throughput analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting
10 library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique that enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify NOVX variants. *See, e.g., Arkin and Yourvan, 1992. Proc. Natl. Acad. Sci. USA 89: 7811-7815; Delgrave, et al., 1993. Protein Engineering 6:327-331.*
15

NOVX Antibodies

The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, i.e., molecules that
20 contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , F_{ab}' and $F_{(ab)2}$ fragments, and an F_{ab} expression library. In general, antibody molecules obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule.
25 Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated protein of the invention intended to serve as an antigen, or a portion or
30 fragment thereof, can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the

invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, such as an amino acid sequence shown in SEQ ID NO: 2n, wherein n is an integer between 1 and 101, and encompasses an epitope thereof such
5 that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues, or at least 30 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its
10 surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of NOVX that is located on the surface of the protein, *e.g.*, a hydrophilic region. A hydrophobicity analysis of the human NOVX protein sequence will indicate which regions of a NOVX polypeptide are particularly hydrophilic and, therefore,
15 are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, *e.g.*, Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-
20 3828; Kyte and Doolittle 1982, *J. Mol. Biol.* 157: 105-142, each incorporated herein by reference in their entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog
25 thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example,
30 Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

Polyclonal Antibodies

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

Monoclonal Antibodies

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one

molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also

have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

5 The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal
10 antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980). It is an objective, especially important in therapeutic applications of monoclonal antibodies, to identify antibodies having a high degree of specificity and a high binding affinity for the target antigen.

After the desired hybridoma cells are identified, the clones can be subcloned by
15 limiting dilution procedures and grown by standard methods (Goding, 1986). Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the
20 culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal
25 antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian
30 COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for example, by substituting the

coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or
5 can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

Humanized Antibodies

The antibodies directed against the protein antigens of the invention can further
10 comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) that are principally comprised
15 of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human
20 antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two,
25 variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and
30 Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

Human Antibodies

Fully human antibodies essentially relate to antibody molecules in which the entire sequence of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein. Human monoclonal antibodies can be prepared by the trioma
5 technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 *Immunol Today* 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp: 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al.,
10 1983. *Proc Natl Acad Sci USA* 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991);
15 Marks et al., *J. Mol. Biol.*, 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire.
20 This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (*Bio/Technology* 10, 779-783 (1992)); Lonberg et al. (*Nature* 368 856-859 (1994)); Morrison (*Nature* 368, 812-13 (1994)); Fishwild et al. (*Nature Biotechnology* 14, 845-51 (1996)); Neuberger (*Nature Biotechnology* 14, 826 (1996)); and Lonberg and Huszar (*Intern. Rev. Immunol.*
25 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light
30 immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial

chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the

- 5 XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies.
- 10 Additionally, the genes encoding the immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

- An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S.
- 15 Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem
- 20 cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

- A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in
- 25 culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

- In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that
- 30 binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

F_{ab} Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an F_{(ab)₂} fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an F_{(ab)₂} fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. $F(ab')_2$ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate $F(ab')_2$ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab' -TNB derivatives is then reconverted to the Fab' -thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab' -TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., J. Immunol. 147:60 (1991).

Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the

particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen
5 described herein and further binds tissue factor (TF).

Heteroconjugate Antibodies

Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such
10 antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide
15 exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector
20 function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992).
25 Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).
30

Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

5 Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, *Phytolaca americana* proteins
10 (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

Conjugates of the antibody and cytotoxic agent are made using a variety of
15 bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-
20 ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See
25 WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such as streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is
30 in turn conjugated to a cytotoxic agent.

Immunoliposomes

The antibodies disclosed herein can also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545.

5 Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the
10 antibody of the present invention can be conjugated to the liposomes as described in Martin et al., J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon et al., J. National Cancer Inst., 81(19): 1484 (1989).

Diagnostic Applications of Antibodies Directed Against the Proteins of the Invention

15 Antibodies directed against a protein of the invention may be used in methods known within the art relating to the localization and/or quantitation of the protein (e.g., for use in measuring levels of the protein within appropriate physiological samples, for use in diagnostic methods, for use in imaging the protein, and the like). In a given embodiment, antibodies against the proteins, or derivatives, fragments, analogs or homologs thereof,
20 that contain the antigen binding domain, are utilized as pharmacologically-active compounds (see below).

An antibody specific for a protein of the invention can be used to isolate the protein by standard techniques, such as immunoaffinity chromatography or immunoprecipitation. Such an antibody can facilitate the purification of the natural
25 protein antigen from cells and of recombinantly produced antigen expressed in host cells. Moreover, such an antibody can be used to detect the antigenic protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the antigenic protein. Antibodies directed against the protein can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example,
30 determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (i.e., physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials,

luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include
5 umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I , ^{131}I , ^{35}S or ^3H .

Antibody Therapeutics

10 Antibodies of the invention, including polyclonal, monoclonal, humanized and fully human antibodies, may be used as therapeutic agents. Such agents will generally be employed to treat or prevent a disease or pathology in a subject. An antibody preparation, preferably one having high specificity and high affinity for its target antigen, is administered to the subject and will generally have an effect due to its binding with the
15 target. Such an effect may be one of two kinds, depending on the specific nature of the interaction between the given antibody molecule and the target antigen in question. In the first instance, administration of the antibody may abrogate or inhibit the binding of the target with an endogenous ligand to which it naturally binds. In this case, the antibody binds to the target and masks a binding site of the naturally occurring ligand, wherein the
20 ligand serves as an effector molecule. Thus the receptor mediates a signal transduction pathway for which ligand is responsible.

 Alternatively, the effect may be one in which the antibody elicits a physiological result by virtue of binding to an effector binding site on the target molecule. In this case the target, a receptor having an endogenous ligand which may be absent or defective in the
25 disease or pathology, binds the antibody as a surrogate effector ligand, initiating a receptor-based signal transduction event by the receptor.

 A therapeutically effective amount of an antibody of the invention relates generally to the amount needed to achieve a therapeutic objective. As noted above, this may be a binding interaction between the antibody and its target antigen that, in certain cases,
30 interferes with the functioning of the target, and in other cases, promotes a physiological response. The amount required to be administered will furthermore depend on the binding affinity of the antibody for its specific antigen, and will also depend on the rate at which

an administered antibody is depleted from the free volume other subject to which it is administered. Common ranges for therapeutically effective dosing of an antibody or antibody fragment of the invention may be, by way of nonlimiting example, from about 0.1 mg/kg body weight to about 50 mg/kg body weight. Common dosing frequencies may range, for example, from twice daily to once a week.

Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a protein of the invention, as well as other molecules identified by the screening assays disclosed herein, can be administered for the treatment of various disorders in the form of pharmaceutical compositions. Principles and considerations involved in preparing such compositions, as well as guidance in the choice of components are provided, for example, in Remington : The Science And Practice Of Pharmacy 19th ed. (Alfonso R. Gennaro, et al., editors) Mack Pub. Co., Easton, Pa. : 1995; Drug Absorption Enhancement : Concepts, Possibilities, Limitations, And Trends, Harwood Academic Publishers, Langhorne, Pa., 1994; and Peptide And Protein Drug Delivery (Advances In Parenteral Sciences, Vol. 4), 1991, M. Dekker, New York.

If the antigenic protein is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein can also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition can comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients can also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example,

hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions.

5 The formulations to be used for in vivo administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations can be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, or
10 microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOTTM (injectable microspheres composed of lactic acid-glycolic acid
15 copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods.

ELISA Assay

An agent for detecting an analyte protein is an antibody capable of binding to an
20 analyte protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., F_{ab} or F_{(ab)2}) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect
25 labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject,
30 as well as tissues, cells and fluids present within a subject. Included within the usage of the term "biological sample", therefore, is blood and a fraction or component of blood including blood serum, blood plasma, or lymph. That is, the detection method of the

invention can be used to detect an analyte mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of an analyte mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of an analyte protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of an analyte genomic DNA include Southern hybridizations. Procedures for conducting immunoassays are described, for example in "ELISA: Theory and Practice: Methods in Molecular Biology", Vol. 42, J. R. Crowther (Ed.) Human Press, Totowa, NJ, 1995; "Immunoassay", E. Diamandis and T. Christopoulos, Academic Press, Inc., San Diego, CA, 1996; and "Practice and Theory of Enzyme Immunoassays", P. Tijssen, Elsevier Science Publishers, Amsterdam, 1985. Furthermore, *in vivo* techniques for detection of an analyte protein include introducing into a subject a labeled anti-analyte protein antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

NOVX Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an NOVX protein, or derivatives, fragments, analogs or homologs thereof. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively-linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and

"vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

5 The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector,
10 "operably-linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner that allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell).

 The term "regulatory sequence" is intended to include promoters, enhancers and
15 other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell and those that direct expression of the nucleotide sequence only in
20 certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic
25 acids as described herein (*e.g.*, NOVX proteins, mutant forms of NOVX proteins, fusion proteins, etc.).

 The recombinant expression vectors of the invention can be designed for expression of NOVX proteins in prokaryotic or eukaryotic cells. For example, NOVX proteins can be expressed in bacterial cells such as *Escherichia coli*, insect cells (using
30 baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, GENE EXPRESSION TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990). Alternatively, the

recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *Escherichia coli* with vectors containing constitutive or inducible promoters directing the expression of either

5 fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: (i) to increase expression of recombinant protein; (ii) to increase the solubility of the recombinant protein; and (iii) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion
10 expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson,
15 1988. *Gene* 67: 31-40), pMAL (New England Biolabs, Beverly, Mass.) and pRIT5 (Pharmacia, Piscataway, N.J.) that fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amrann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, GENE EXPRESSION
20 TECHNOLOGY: METHODS IN ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 60-89).

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein. *See, e.g.*, Gottesman, GENE EXPRESSION TECHNOLOGY: METHODS IN
25 ENZYMOLOGY 185, Academic Press, San Diego, Calif. (1990) 119-128. Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (*see, e.g.*, Wada, *et al.*, 1992. *Nucl. Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis
30 techniques.

In another embodiment, the NOVX expression vector is a yeast expression vector. Examples of vectors for expression in yeast *Saccharomyces cerevisiae* include pYepSec1

(Baldari, *et al.*, 1987. *EMBO J.* 6: 229-234), pMFa (Kurjan and Herskowitz, 1982. *Cell* 30: 933-943), pJRY88 (Schultz *et al.*, 1987. *Gene* 54: 113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.), and picZ (Invitrogen Corp, San Diego, Calif.).

Alternatively, NOVX can be expressed in insect cells using baculovirus expression
5 vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., SF9 cells) include the pAc series (Smith, *et al.*, 1983. *Mol. Cell. Biol.* 3: 2156-2165) and the pVL series (Lucklow and Summers, 1989. *Virology* 170: 31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in
mammalian cells using a mammalian expression vector. Examples of mammalian
10 expression vectors include pCDM8 (Seed, 1987. *Nature* 329: 840) and pMT2PC (Kaufman, *et al.*, 1987. *EMBO J.* 6: 187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, adenovirus 2, cytomegalovirus, and simian virus 40. For other suitable expression systems for both
15 prokaryotic and eukaryotic cells see, e.g., Chapters 16 and 17 of Sambrook, *et al.*, MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

In another embodiment, the recombinant mammalian expression vector is capable
of directing expression of the nucleic acid preferentially in a particular cell type (e.g.,
20 tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert, *et al.*, 1987. *Genes Dev.* 1: 268-277), lymphoid-specific promoters (Calame and Eaton, 1988. *Adv. Immunol.* 43: 235-275), in particular promoters of T cell receptors (Winoto and
25 Baltimore, 1989. *EMBO J.* 8: 729-733) and immunoglobulins (Banerji, *et al.*, 1983. *Cell* 33: 729-740; Queen and Baltimore, 1983. *Cell* 33: 741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle, 1989. *Proc. Natl. Acad. Sci. USA* 86: 5473-5477), pancreas-specific promoters (Edlund, *et al.*, 1985. *Science* 230: 912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Pat. No. 4,873,316
30 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, e.g., the murine hox promoters (Kessel and Gruss, 1990).

Science 249: 374-379) and the α -fetoprotein promoter (Campes and Tilghman, 1989. *Genes Dev.* 3: 537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation.

- 5 That is, the DNA molecule is operatively-linked to a regulatory sequence in a manner that allows for expression (by transcription of the DNA molecule) of an RNA molecule that is antisense to NOVX mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen that direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or
- 10 enhancers, or regulatory sequences can be chosen that direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced.
- 15 For a discussion of the regulation of gene expression using antisense genes *see, e.g.,* Weintraub, *et al.*, "Antisense RNA as a molecular tool for genetic analysis," *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

- Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and
- 20 "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but also to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

- 25 A host cell can be any prokaryotic or eukaryotic cell. For example, NOVX protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

- Vector DNA can be introduced into prokaryotic or eukaryotic cells via
- 30 conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including

calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (MOLECULAR CLONING: A LABORATORY MANUAL. 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Various selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding NOVX or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) NOVX protein. Accordingly, the invention further provides methods for producing NOVX protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding NOVX protein has been introduced) in a suitable medium such that NOVX protein is produced. In another embodiment, the method further comprises isolating NOVX protein from the medium or the host cell.

25 Transgenic NOVX Animals

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which NOVX protein-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous NOVX sequences have been introduced into their genome or homologous recombinant animals in which endogenous NOVX sequences have been altered. Such animals are useful for studying the function and/or activity of NOVX

protein and for identifying and/or evaluating modulators of NOVX protein activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA that is integrated into the genome of a cell from which a transgenic animal develops and that remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous NOVX gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing NOVX-encoding nucleic acid into the male pronuclei of a fertilized oocyte (*e.g.*, by microinjection, retroviral infection) and allowing the oocyte to develop in a pseudopregnant female foster animal. The human NOVX cDNA sequences SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a non-human homologue of the human NOVX gene, such as a mouse NOVX gene, can be isolated based on hybridization to the human NOVX cDNA (described further *supra*) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably-linked to the NOVX transgene to direct expression of NOVX protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866; 4,870,009; and 4,873,191; and Hogan, 1986. In: MANIPULATING THE MOUSE EMBRYO, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the NOVX transgene in its genome and/or expression of NOVX mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed

additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene-encoding NOVX protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains
5 at least a portion of an NOVX gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the NOVX gene. The NOVX gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101), but more preferably, is a non-human homologue of a human NOVX gene. For example, a mouse homologue of human NOVX gene of SEQ ID NO: 2n-1,
10 wherein n is an integer between 1 and 101 can be used to construct a homologous recombination vector suitable for altering an endogenous NOVX gene in the mouse genome. In one embodiment, the vector is designed such that, upon homologous recombination, the endogenous NOVX gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector).

15 Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous NOVX gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous NOVX protein). In the homologous recombination vector, the altered portion of the NOVX gene is flanked at its 5'- and 3'-termini by
20 additional nucleic acid of the NOVX gene to allow for homologous recombination to occur between the exogenous NOVX gene carried by the vector and an endogenous NOVX gene in an embryonic stem cell. The additional flanking NOVX nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5'- and 3'-termini) are included
25 in the vector. *See, e.g.*, Thomas, *et al.*, 1987. *Cell* 51: 503 for a description of homologous recombination vectors. The vector is then introduced into an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced NOVX gene has homologously-recombined with the endogenous NOVX gene are selected. *See, e.g.*, Li, *et al.*, 1992. *Cell* 69: 915.

30 The selected cells are then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras. *See, e.g.*, Bradley, 1987. In: TERATOCARCINOMAS AND EMBRYONIC STEM CELLS: A PRACTICAL APPROACH, Robertson, ed. IRL, Oxford, pp.

113-152. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously-recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously-recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley, 1991. *Curr. Opin. Biotechnol.* 2: 823-829; PCT International Publication Nos.: WO 90/11354; WO 91/01140; WO 92/0968; and WO 93/04169.

In another embodiment, transgenic non-humans animals can be produced that contain selected systems that allow for regulated expression of the transgene. One example of such a system is the cre/loxP recombinase system of bacteriophage P1. For a description of the cre/loxP recombinase system, See, e.g., Lakso, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae*. See, O'Gorman, *et al.*, 1991. *Science* 251:1351-1355. If a cre/loxP recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the Cre recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, *et al.*, 1997. *Nature* 385: 810-813. In brief, a cell (e.g., a somatic cell) from the transgenic animal can be isolated and induced to exit the growth cycle and enter G₀ phase. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell (e.g., the somatic cell) is isolated.

30 **Pharmaceutical Compositions**

The NOVX nucleic acid molecules, NOVX proteins, and anti-NOVX antibodies (also referred to herein as "active compounds") of the invention, and derivatives,

fragments, analogs and homologs thereof, can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein, "pharmaceutically acceptable carrier" is intended to include any and all
5 solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Suitable carriers are described in the most recent edition of Remington's Pharmaceutical Sciences, a standard reference text in the field, which is incorporated herein by reference. Preferred examples of such carriers or diluents include, but are not limited to, water,
10 saline, finger's solutions, dextrose solution, and 5% human serum albumin. Liposomes and non-aqueous vehicles such as fixed oils may also be used. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be
15 incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (*i.e.*, topical), transmucosal, and rectal administration. Solutions or suspensions used for
20 parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid (EDTA);
25 buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous
30 solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™

(BASF, Parsippany, N.J.) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringeability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi.

5 The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of
10 microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent
15 which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., an NOVX protein or anti-NOVX antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active
20 compound into a sterile vehicle that contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, methods of preparation are vacuum drying and freeze-drying that yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

25 Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is
30 applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or

compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used

herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and
5 directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for
10 example, intravenous injection, local administration (*see, e.g.*, U.S. Patent No. 5,328,470) or by stereotactic injection (*see, e.g.*, Chen, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene
15 delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells that produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

20 Screening and Detection Methods

The isolated nucleic acid molecules of the invention can be used to express NOVX protein (*e.g.*, via a recombinant expression vector in a host cell in gene therapy applications), to detect NOVX mRNA (*e.g.*, in a biological sample) or a genetic lesion in an NOVX gene, and to modulate NOVX activity, as described further, below. In addition,
25 the NOVX proteins can be used to screen drugs or compounds that modulate the NOVX protein activity or expression as well as to treat disorders characterized by insufficient or excessive production of NOVX protein or production of NOVX protein forms that have decreased or aberrant activity compared to NOVX wild-type protein (*e.g.*; diabetes (regulates insulin release); obesity (binds and transport lipids); metabolic disturbances
30 associated with obesity, the metabolic syndrome X as well as anorexia and wasting disorders associated with chronic diseases and various cancers, and infectious disease (possesses anti-microbial activity) and the various dyslipidemias. In addition, the

anti-NOVX antibodies of the invention can be used to detect and isolate NOVX proteins and modulate NOVX activity. In yet a further aspect, the invention can be used in methods to influence appetite, absorption of nutrients and the disposition of metabolic substrates in both a positive and negative fashion.

- 5 The invention further pertains to novel agents identified by the screening assays described herein and uses thereof for treatments as described, *supra*.

Screening Assays

- The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, 10 peptidomimetics, small molecules or other drugs) that bind to NOVX proteins or have a stimulatory or inhibitory effect on, *e.g.*, NOVX protein expression or NOVX protein activity. The invention also includes compounds identified in the screening assays described herein.

- In one embodiment, the invention provides assays for screening candidate or test 15 compounds which bind to or modulate the activity of the membrane-bound form of an NOVX protein or polypeptide or biologically-active portion thereof. The test compounds of the invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring 20 deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds. *See, e.g.*, Lam, 1997. *Anticancer Drug Design* 12: 145.

- 25 A "small molecule" as used herein, is meant to refer to a composition that has a molecular weight of less than about 5 kD and most preferably less than about 4 kD. Small molecules can be, *e.g.*, nucleic acids, peptides, polypeptides, peptidomimetics, carbohydrates, lipids or other organic or inorganic molecules. Libraries of chemical and/or biological mixtures, such as fungal, bacterial, or algal extracts, are known in the art 30 and can be screened with any of the assays of the invention.

 Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt, *et al.*, 1993. *Proc. Natl. Acad. Sci. U.S.A.* 90: 6909; Erb, *et al.*,

1994. *Proc. Natl. Acad. Sci. U.S.A.* 91: 11422; Zuckermann, *et al.*, 1994. *J. Med. Chem.* 37: 2678; Cho, *et al.*, 1993. *Science* 261: 1303; Carrell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2059; Carell, *et al.*, 1994. *Angew. Chem. Int. Ed. Engl.* 33: 2061; and Gallop, *et al.*, 1994. *J. Med. Chem.* 37: 1233.

- 5 Libraries of compounds may be presented in solution (*e.g.*, Houghten, 1992. *Biotechniques* 13: 412-421), or on beads (Lam, 1991. *Nature* 354: 82-84), on chips (Fodor, 1993. *Nature* 364: 555-556), bacteria (Ladner, U.S. Patent No. 5,223,409), spores (Ladner, U.S. Patent 5,233,409), plasmids (Cull, *et al.*, 1992. *Proc. Natl. Acad. Sci. USA* 89: 1865-1869) or on phage (Scott and Smith, 1990. *Science* 249: 386-390; Devlin, 1990. *Science* 249: 404-406; Cwirla, *et al.*, 1990. *Proc. Natl. Acad. Sci. U.S.A.* 87: 6378-6382; Felici, 1991. *J. Mol. Biol.* 222: 301-310; Ladner, U.S. Patent No. 5,233,409.).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to an NOVX protein determined. The cell, for example, can of mammalian origin or a yeast cell. Determining the ability of the test compound to bind to the NOVX protein can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the NOVX protein or biologically-active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I , ^{35}S , ^{14}C , or ^3H , either directly or indirectly, and the radioisotope detected by direct counting of radioemission or by scintillation counting. Alternatively, test compounds can be enzymatically-labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In one embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX protein or a biologically-active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of NOVX protein, or a biologically-active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the NOVX protein or biologically-active portion thereof. Determining the ability of the test compound to modulate the activity of NOVX or a biologically-active portion thereof can be accomplished, for example, by determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule. As used herein, a "target molecule" is a molecule with which an NOVX protein binds or interacts in nature, for example, a molecule on the surface of a cell which expresses an NOVX interacting protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. An NOVX target molecule can be a non-NOVX molecule or an NOVX protein or polypeptide of the invention. In one embodiment, an NOVX target molecule is a component of a signal transduction pathway that facilitates transduction of an extracellular signal (*e.g.* a signal generated by binding of a compound to a membrane-bound NOVX molecule) through the cell membrane and into the cell. The target, for example, can be a second intercellular protein that has catalytic activity or a protein that facilitates the association of downstream signaling molecules with NOVX.

Determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule can be accomplished by one of the methods described above for determining direct binding. In one embodiment, determining the ability of the NOVX protein to bind to or interact with an NOVX target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.* intracellular Ca^{2+} , diacylglycerol, IP_3 , etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (comprising an NOVX-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a cellular response, for example, cell survival, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the invention is a cell-free assay comprising contacting an NOVX protein or biologically-active portion thereof with a test

compound and determining the ability of the test compound to bind to the NOVX protein or biologically-active portion thereof. Binding of the test compound to the NOVX protein can be determined either directly or indirectly as described above. In one such embodiment, the assay comprises contacting the NOVX protein or biologically-active
5 portion thereof with a known compound which binds NOVX to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the test compound to preferentially bind to NOVX or biologically-active portion thereof as
10 compared to the known compound.

In still another embodiment, an assay is a cell-free assay comprising contacting NOVX protein or biologically-active portion thereof with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the NOVX protein or biologically-active portion thereof. Determining the
15 ability of the test compound to modulate the activity of NOVX can be accomplished, for example, by determining the ability of the NOVX protein to bind to an NOVX target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of NOVX protein can be accomplished by determining the ability of the NOVX
20 protein further modulate an NOVX target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as described, *supra*.

In yet another embodiment, the cell-free assay comprises contacting the NOVX protein or biologically-active portion thereof with a known compound which binds NOVX
25 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with an NOVX protein, wherein determining the ability of the test compound to interact with an NOVX protein comprises determining the ability of the NOVX protein to preferentially bind to or modulate the activity of an NOVX target molecule.

30 The cell-free assays of the invention are amenable to use of both the soluble form or the membrane-bound form of NOVX protein. In the case of cell-free assays comprising the membrane-bound form of NOVX protein, it may be desirable to utilize a solubilizing

agent such that the membrane-bound form of NOVX protein is maintained in solution.

Examples of such solubilizing agents include non-ionic detergents such as

n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton[®] X-100, Triton[®] X-114, Thesit[®],

- 5 Isotridecypoly(ethylene glycol ether)_n, N-dodecyl--N,N-dimethyl-3-ammonio-1-propane sulfonate, 3-(3-cholamidopropyl) dimethylamminiol-1-propane sulfonate (CHAPS), or 3-(3-cholamidopropyl)dimethylamminiol-2-hydroxy-1-propane sulfonate (CHAPSO).

- In more than one embodiment of the above assay methods of the invention, it may be desirable to immobilize either NOVX protein or its target molecule to facilitate
- 10 separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to NOVX protein, or interaction of NOVX protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes,
- 15 and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided that adds a domain that allows one or both of the proteins to be bound to a matrix. For example, GST-NOVX fusion proteins or GST-target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, that are then combined with the test compound or the test compound and
- 20 either the non-adsorbed target protein or NOVX protein, and the mixture is incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described, *supra*. Alternatively,
- 25 the complexes can be dissociated from the matrix, and the level of NOVX protein binding or activity determined using standard techniques.

- Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the NOVX protein or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated
- 30 NOVX protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well-known within the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, Ill.), and immobilized in the wells of streptavidin-coated

96 well plates (Pierce Chemical). Alternatively, antibodies reactive with NOVX protein or target molecules, but which do not interfere with binding of the NOVX protein to its target molecule, can be derivatized to the wells of the plate, and unbound target or NOVX protein trapped in the wells by antibody conjugation. Methods for detecting such
5 complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the NOVX protein or target molecule, as well as enzyme-linked assays that rely on detecting an enzymatic activity associated with the NOVX protein or target molecule.

In another embodiment, modulators of NOVX protein expression are identified in
10 a method wherein a cell is contacted with a candidate compound and the expression of NOVX mRNA or protein in the cell is determined. The level of expression of NOVX mRNA or protein in the presence of the candidate compound is compared to the level of expression of NOVX mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of NOVX mRNA or protein
15 expression based upon this comparison. For example, when expression of NOVX mRNA or protein is greater (*i.e.*, statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of NOVX mRNA or protein expression. Alternatively, when expression of NOVX mRNA or protein is less (statistically significantly less) in the presence of the candidate compound
20 than in its absence, the candidate compound is identified as an inhibitor of NOVX mRNA or protein expression. The level of NOVX mRNA or protein expression in the cells can be determined by methods described herein for detecting NOVX mRNA or protein.

In yet another aspect of the invention, the NOVX proteins can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317;
25 Zervos, *et al.*, 1993. *Cell* 72: 223-232; Madura, *et al.*, 1993. *J. Biol. Chem.* 268: 12046-12054; Bartel, *et al.*, 1993. *Biotechniques* 14: 920-924; Iwabuchi, *et al.*, 1993. *Oncogene* 8: 1693-1696; and Brent WO 94/10300), to identify other proteins that bind to or interact with NOVX ("NOVX-binding proteins" or "NOVX-bp") and modulate NOVX activity. Such NOVX-binding proteins are also likely to be involved in the propagation of
30 signals by the NOVX proteins as, for example, upstream or downstream elements of the NOVX pathway.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for NOVX is fused to a gene encoding the DNA binding domain of a known transcription factor (e.g., GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are able to interact, *in vivo*, forming an NOVX-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (e.g., LacZ) that is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene that encodes the protein which interacts with NOVX.

The invention further pertains to novel agents identified by the aforementioned screening assays and uses thereof for treatments as described herein.

Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. By way of example, and not of limitation, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. Some of these applications are described in the subsections, below.

Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the NOVX sequences, SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or fragments or derivatives thereof, can be used to map the location of the NOVX genes, respectively, on a

chromosome. The mapping of the NOVX sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, NOVX genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the NOVX sequences. Computer analysis of the NOVX, sequences can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the NOVX sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (e.g., human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but in which human cells can, the one human chromosome that contains the gene encoding the needed enzyme will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. See, e.g., D'Eustachio, *et al.*, 1983. *Science* 220: 919-924. Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the NOVX sequences to design oligonucleotide primers, sub-localization can be achieved with panels of fragments from specific chromosomes.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical like colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be

identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases; and more preferably 2,000 bases, will suffice to get
5 good results at a reasonable amount of time. For a review of this technique, *see*, Verma, *et al.*, HUMAN CHROMOSOMES: A MANUAL OF BASIC TECHNIQUES (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for
10 marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical
15 position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, *e.g.*, in McKusick, MENDELIAN INHERITANCE IN MAN, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described
20 in, *e.g.*, Egeland, *et al.*, 1987. *Nature*, 325: 783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the NOVX gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease.
25 Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

30 Tissue Typing

The NOVX sequences of the invention can also be used to identify individuals from minute biological samples. In this technique, an individual's genomic DNA is

digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. The sequences of the invention are useful as additional DNA markers for RFLP ("restriction fragment length polymorphisms," described in U.S. Patent No. 5,272,057).

5 Furthermore, the sequences of the invention can be used to provide an alternative technique that determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the NOVX sequences described herein can be used to prepare two PCR primers from the 5'- and 3'-termini of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

10 Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the invention can be used to obtain such identification sequences from individuals and from tissue. The NOVX sequences of the invention uniquely represent portions of the human
15 genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Much of the allelic variation is due to single nucleotide polymorphisms (SNPs), which include restriction fragment length polymorphisms (RFLPs).

20 Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000
25 primers that each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

Predictive Medicine

30 The invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically.

Accordingly, one aspect of the invention relates to diagnostic assays for determining NOVX protein and/or nucleic acid expression as well as NOVX activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant NOVX expression or activity. The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with NOVX protein, nucleic acid expression or activity. For example, mutations in an NOVX gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with NOVX protein, nucleic acid expression, or biological activity.

Another aspect of the invention provides methods for determining NOVX protein, nucleic acid expression or activity in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (*e.g.*, drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (*e.g.*, the genotype of the individual examined to determine the ability of the individual to respond to a particular agent.) Yet another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX in clinical trials.

These and other agents are described in further detail in the following sections.

Diagnostic Assays

An exemplary method for detecting the presence or absence of NOVX in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes NOVX protein such that the presence of NOVX is detected in the biological sample. An agent for detecting NOVX

mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to NOVX mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length NOVX nucleic acid, such as the nucleic acid of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to NOVX mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

An agent for detecting NOVX protein is an antibody capable of binding to NOVX protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or F(ab')₂) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect NOVX mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of NOVX mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of NOVX protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations, and immunofluorescence. *In vitro* techniques for detection of NOVX genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of NOVX protein include introducing into a subject a labeled anti-NOVX antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting NOVX protein, mRNA, or genomic DNA, such that the presence of NOVX protein, mRNA or genomic DNA is detected in the biological sample, and
5 comparing the presence of NOVX protein, mRNA or genomic DNA in the control sample with the presence of NOVX protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of NOVX in a biological sample. For example, the kit can comprise: a labeled compound or agent capable of detecting NOVX protein or mRNA in a biological sample; means for
10 determining the amount of NOVX in the sample; and means for comparing the amount of NOVX in the sample with a standard. The compound or agent can be packaged in a suitable container. The kit can further comprise instructions for using the kit to detect NOVX protein or nucleic acid.

15 Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. For example, the assays described herein, such as the
20 preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with NOVX protein, nucleic acid expression or activity. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disease or disorder. Thus, the invention provides a method for identifying a disease or disorder associated with aberrant NOVX expression
25 or activity in which a test sample is obtained from a subject and NOVX protein or nucleic acid (*e.g.*, mRNA, genomic DNA) is detected, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant NOVX expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test
30 sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to

treat a disease or disorder associated with aberrant NOVX expression or activity. For example, such methods can be used to determine whether a subject can be effectively treated with an agent for a disorder. Thus, the invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with
5 aberrant NOVX expression or activity in which a test sample is obtained and NOVX protein or nucleic acid is detected (*e.g.*, wherein the presence of NOVX protein or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant NOVX expression or activity).

The methods of the invention can also be used to detect genetic lesions in an
10 NOVX gene, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In various embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of an alteration affecting the integrity of a gene encoding an NOVX-protein, or the misexpression of the
15 NOVX gene. For example, such genetic lesions can be detected by ascertaining the existence of at least one of: (i) a deletion of one or more nucleotides from an NOVX gene; (ii) an addition of one or more nucleotides to an NOVX gene; (iii) a substitution of one or more nucleotides of an NOVX gene, (iv) a chromosomal rearrangement of an NOVX gene; (v) an alteration in the level of a messenger RNA transcript of an NOVX gene, (vi)
20 aberrant modification of an NOVX gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild-type splicing pattern of a messenger RNA transcript of an NOVX gene, (viii) a non-wild-type level of an NOVX protein, (ix) allelic loss of an NOVX gene, and (x) inappropriate post-translational modification of an NOVX protein. As described herein, there are a large number of assay techniques known in the
25 art which can be used for detecting lesions in an NOVX gene. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject. However, any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

In certain embodiments, detection of the lesion involves the use of a probe/primer
30 in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran, *et al.*, 1988. *Science* 241: 1077-1080; and

Nakazawa, *et al.*, 1994. *Proc. Natl. Acad. Sci. USA* 91: 360-364), the latter of which can be particularly useful for detecting point mutations in the NOVX-gene (*see*, Abravaya, *et al.*, 1995. *Nucl. Acids Res.* 23: 675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers that specifically hybridize to an NOVX gene under conditions such that hybridization and amplification of the NOVX gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (*see*, Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), transcriptional amplification system (*see*, Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 1173-1177); Q β Replicase (*see*, Lizardi, *et al.*, 1988. *BioTechnology* 6: 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in an NOVX gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.*, U.S. Patent No. 5,493,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in NOVX can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high-density arrays containing hundreds or thousands of oligonucleotides probes. *See, e.g.*, Cronin, *et al.*, 1996. *Human Mutation* 7: 244-255; Kozal, *et al.*, 1996. *Nat. Med.* 2: 753-759. For example, genetic mutations in NOVX can be identified in two dimensional arrays

containing light-generated DNA probes as described in Cronin, *et al.*, *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the NOVX gene and detect mutations by comparing the sequence of the sample NOVX with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert, 1977. *Proc. Natl. Acad. Sci. USA* 74: 560 or Sanger, 1977. *Proc. Natl. Acad. Sci. USA* 74: 5463. It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays (*see, e.g.*, Naeve, *et al.*, 1995. *Biotechniques* 19: 448), including sequencing by mass spectrometry (*see, e.g.*, PCT International Publication No. WO 94/16101; Cohen, *et al.*, 1996. *Adv. Chromatography* 36: 127-162; and Griffin, *et al.*, 1993. *Appl. Biochem. Biotechnol.* 38: 147-159).

Other methods for detecting mutations in the NOVX gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes. *See, e.g.*, Myers, *et al.*, 1985. *Science* 230: 1242. In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type NOVX sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent that cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S₁ nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then

separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, e.g., Cotton, *et al.*, 1988. *Proc. Natl. Acad. Sci. USA* 85: 4397; Saleeba, *et al.*, 1992. *Methods Enzymol.* 217: 286-295. In an embodiment, the control DNA or RNA can be labeled for detection.

5 In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in NOVX cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells
10 cleaves T at G/T mismatches. See, e.g., Hsu, *et al.*, 1994. *Carcinogenesis* 15: 1657-1662. According to an exemplary embodiment, a probe based on an NOVX sequence, e.g., a wild-type NOVX sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, e.g., U.S.
15 Patent No. 5,459,039.

 In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in NOVX genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids. See, e.g., Orita, *et al.*, 1989. *Proc. Natl.*
20 *Acad. Sci. USA*: 86: 2766; Cotton, 1993. *Mutat. Res.* 285: 125-144; Hayashi, 1992. *Genet. Anal. Tech. Appl.* 9: 73-79. Single-stranded DNA fragments of sample and control NOVX nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA
25 fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In one embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility. See, e.g., Keen, *et al.*, 1991. *Trends Genet.* 7: 5.

30 In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE). See, e.g., Myers, *et al.*, 1985. *Nature* 313: 495.

When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA. See, e.g., Rosenbaum and Reissner, 1987. *Biophys. Chem.* 265: 12753.

5 Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions that permit hybridization only if a perfect match is found. See, e.g., Saiki, *et al.*, 1986. *Nature* 324: 163; Saiki, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* 86: 6230. Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

15 Alternatively, allele specific amplification technology that depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization; see, e.g., Gibbs, *et al.*, 1989. *Nucl. Acids Res.* 17: 2437-2448) or at the extreme 3'-terminus of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (see, e.g., Prossner, 1993. *Tibtech.* 11: 238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection. See, e.g., Gasparini, *et al.*, 1992. *Mol. Cell Probes* 6: 1. It is anticipated that in certain embodiments amplification may also be performed using *Taq* ligase for amplification. See, e.g., Barany, 1991. *Proc. Natl. Acad. Sci. USA* 88: 189. In such cases, ligation will occur only if there is a perfect match at the 3'-terminus of the 5' sequence, making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

25 The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to

diagnose patients exhibiting symptoms or family history of a disease or illness involving an NOVX gene.

Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which NOVX is expressed may be utilized in the prognostic assays described herein. However,
5 any biological sample containing nucleated cells may be used, including, for example, buccal mucosal cells.

Pharmacogenomics

10 Agents, or modulators that have a stimulatory or inhibitory effect on NOVX activity (*e.g.*, NOVX gene expression), as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders (The disorders include metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease,
15 Parkinson's Disorder, immune disorders, and hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.) In conjunction with such treatment, the pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or
20 drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (*e.g.*, drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such
25 pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of NOVX protein, expression of NOVX nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the
30 response to drugs due to altered drug disposition and abnormal action in affected persons. See *e.g.*, Eichelbaum, 1996. *Clin. Exp. Pharmacol. Physiol.*, 23: 983-985; Linder, 1997. *Clin. Chem.*, 43: 254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act

on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase (G6PD) deficiency is a common inherited
5 enzymopathy in which the main clinical complication is hemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of
10 genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome PREGNANCY ZONE PROTEIN PRECURSOR enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two
15 phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects
20 when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. At the other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene
25 amplification.

Thus, the activity of NOVX protein, expression of NOVX nucleic acid, or mutation content of NOVX genes in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding
30 drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency

when treating a subject with an NOVX modulator, such as a modulator identified by one of the exemplary screening assays described herein.

Monitoring of Effects During Clinical Trials

5 Monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of NOVX (*e.g.*, the ability to modulate aberrant cell proliferation and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described
10 herein to increase NOVX gene expression, protein levels, or upregulate NOVX activity, can be monitored in clinical trials of subjects exhibiting decreased NOVX gene expression, protein levels, or downregulated NOVX activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease NOVX gene expression, protein levels, or downregulate NOVX activity, can be monitored in clinical
15 trials of subjects exhibiting increased NOVX gene expression, protein levels, or upregulated NOVX activity. In such clinical trials, the expression or activity of NOVX and, preferably, other genes that have been implicated in, for example, a cellular proliferation or immune disorder can be used as a "read out" or markers of the immune responsiveness of a particular cell.

20 By way of example, and not of limitation, genes, including NOVX, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) that modulates NOVX activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the
25 levels of expression of NOVX and other genes implicated in the disorder. The levels of gene expression (*i.e.*, a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of NOVX or other genes. In this manner, the gene expression pattern can serve as
30 a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In one embodiment, the invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, protein, peptide, peptidomimetic, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a
5 pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of an NOVX protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the NOVX protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression
10 or activity of the NOVX protein, mRNA, or genomic DNA in the pre-administration sample with the NOVX protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of NOVX to higher levels than detected, *i.e.*, to increase
15 the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of NOVX to lower levels than detected, *i.e.*, to decrease the effectiveness of the agent.

20 Methods of Treatment

The invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant NOVX expression or activity. The disorders include cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis, atrial septal defect
25 (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia, prostate cancer, neoplasm; adenocarcinoma, lymphoma, uterus cancer, fertility, hemophilia, hypercoagulation, idiopathic thrombocytopenic purpura,
30 immunodeficiencies, graft versus host disease, AIDS, bronchial asthma, Crohn's disease; multiple sclerosis, treatment of Albright Hereditary Osteodystrophy, and other diseases, disorders and conditions of the like.

These methods of treatment will be discussed more fully, below.

Disease and Disorders

Diseases and disorders that are characterized by increased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with
5 Therapeutics that antagonize (*i.e.*, reduce or inhibit) activity. Therapeutics that antagonize activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to: (i) an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; (ii) antibodies to an aforementioned peptide;
10 (iii) nucleic acids encoding an aforementioned peptide; (iv) administration of antisense nucleic acid and nucleic acids that are "dysfunctional" (*i.e.*, due to a heterologous insertion within the coding sequences of coding sequences to an aforementioned peptide) that are utilized to "knockout" endogenous function of an aforementioned peptide by homologous recombination (*see, e.g.*, Capecchi, 1989. *Science* 244: 1288-1292); or (v) modulators (
15 *i.e.*, inhibitors, agonists and antagonists, including additional peptide mimetic of the invention or antibodies specific to a peptide of the invention) that alter the interaction between an aforementioned peptide and its binding partner.

Diseases and disorders that are characterized by decreased (relative to a subject not suffering from the disease or disorder) levels or biological activity may be treated with
20 Therapeutics that increase (*i.e.*, are agonists to) activity. Therapeutics that upregulate activity may be administered in a therapeutic or prophylactic manner. Therapeutics that may be utilized include, but are not limited to, an aforementioned peptide, or analogs, derivatives, fragments or homologs thereof; or an agonist that increases bioavailability.

Increased or decreased levels can be readily detected by quantifying peptide and/or
25 RNA, by obtaining a patient tissue sample (*e.g.*, from biopsy tissue) and assaying it *in vitro* for RNA or peptide levels, structure and/or activity of the expressed peptides (or mRNAs of an aforementioned peptide). Methods that are well-known within the art include, but are not limited to, immunoassays (*e.g.*, by Western blot analysis, immunoprecipitation followed by sodium dodecyl sulfate (SDS) polyacrylamide gel
30 electrophoresis, immunocytochemistry, etc.) and/or hybridization assays to detect expression of mRNAs (*e.g.*, Northern assays, dot blots, *in situ* hybridization, and the like).

Prophylactic Methods

In one aspect, the invention provides a method for preventing, in a subject, a disease or condition associated with an aberrant NOVX expression or activity, by administering to the subject an agent that modulates NOVX expression or at least one NOVX activity. Subjects at risk for a disease that is caused or contributed to by aberrant NOVX expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the NOVX aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending upon the type of NOVX aberrancy, for example, an NOVX agonist or NOVX antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein. The prophylactic methods of the invention are further discussed in the following subsections.

Therapeutic Methods

Another aspect of the invention pertains to methods of modulating NOVX expression or activity for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of NOVX protein activity associated with the cell. An agent that modulates NOVX protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of an NOVX protein, a peptide, an NOVX peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more NOVX protein activity. Examples of such stimulatory agents include active NOVX protein and a nucleic acid molecule encoding NOVX that has been introduced into the cell. In another embodiment, the agent inhibits one or more NOVX protein activity. Examples of such inhibitory agents include antisense NOVX nucleic acid molecules and anti-NOVX antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g., by administering the agent to a subject). As such, the invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of an NOVX protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., up-regulates or down-regulates) NOVX expression or activity. In another embodiment, the method involves administering an NOVX protein or

nucleic acid molecule as therapy to compensate for reduced or aberrant NOVX expression or activity.

Stimulation of NOVX activity is desirable *in situations* in which NOVX is abnormally downregulated and/or in which increased NOVX activity is likely to have a beneficial effect. One example of such a situation is where a subject has a disorder characterized by aberrant cell proliferation and/or differentiation (*e.g.*, cancer or immune associated disorders). Another example of such a situation is where the subject has a gestational disease (*e.g.*, preclampsia).

10 **Determination of the Biological Effect of the Therapeutic**

In various embodiments of the invention, suitable *in vitro* or *in vivo* assays are performed to determine the effect of a specific Therapeutic and whether its administration is indicated for treatment of the affected tissue.

15 In various specific embodiments, *in vitro* assays may be performed with representative cells of the type(s) involved in the patient's disorder, to determine if a given Therapeutic exerts the desired effect upon the cell type(s). Compounds for use in therapy may be tested in suitable animal model systems including, but not limited to rats, mice, chicken, cows, monkeys, rabbits, and the like, prior to testing in human subjects. Similarly, for *in vivo* testing, any of the animal model system known in the art may be used prior to administration to human subjects.

Prophylactic and Therapeutic Uses of the Compositions of the Invention

25 The NOVX nucleic acids and proteins of the invention are useful in potential prophylactic and therapeutic applications implicated in a variety of disorders including, but not limited to: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases and various cancers.

30 As an example, a cDNA encoding the NOVX protein of the invention may be useful in gene therapy, and the protein may be useful when administered to a subject in need thereof. By way of non-limiting example, the compositions of the invention will

have efficacy for treatment of patients suffering from: metabolic disorders, diabetes, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders, hematopoietic disorders, and the various dyslipidemias.

5 Both the novel nucleic acid encoding the NOVX protein, and the NOVX protein of the invention, or fragments thereof, may also be useful in diagnostic applications, wherein the presence or amount of the nucleic acid or the protein are to be assessed. A further use could be as an anti-bacterial molecule (*i.e.*, some peptides have been found to possess anti-bacterial properties). These materials are further useful in the generation of antibodies,
10 which immunospecifically-bind to the novel substances of the invention for use in therapeutic or diagnostic methods.

Sequence Analyses

15 The sequence of NOVX was derived by laboratory cloning of cDNA fragments, by *in silico* prediction of the sequence. cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, were cloned. *In silico* prediction was based on sequences available in CuraGen's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA sequence, or some
20 portion thereof.

The laboratory cloning was performed using one or more of the methods summarized below:

25 **SeqCalling™ Technology:** cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example,
30 growth factors, chemokines or steroids. The cDNA thus derived was then sequenced using CuraGen Corporation's SeqCalling technology which is disclosed in full in U. S. Ser. Nos. 09/417,386 filed Oct. 13, 1999, and 09/614,505 filed July 11, 2000. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all

samples were assembled together, sometimes including public human sequences, using bioinformatics programs to produce a consensus sequence for each assembly. Each assembly is included in CuraGen Corporation's database. Sequences were included as components for assembly when the extent of identity with another component was at least
5 95% over 50 bp. Each assembly represents a gene or portion thereof and includes information on variants, such as splice forms single nucleotide polymorphisms (SNPs), insertions, deletions and other sequence variations.

Variant sequences are also included in this application. A variant sequence can include a single nucleotide polymorphism (SNP). A SNP can, in some instances, be
10 referred to as a "cSNP" to denote that the nucleotide sequence containing the SNP originates as a cDNA. A SNP can arise in several ways. For example, a SNP may be due to a substitution of one nucleotide for another at the polymorphic site. Such a substitution can be either a transition or a transversion. A SNP can also arise from a deletion of a nucleotide or an insertion of a nucleotide, relative to a reference allele. In this case, the
15 polymorphic site is a site at which one allele bears a gap with respect to a particular nucleotide in another allele. SNPs occurring within genes may result in an alteration of the amino acid encoded by the gene at the position of the SNP. Intragenic SNPs may also be silent, when a codon including a SNP encodes the same amino acid as a result of the redundancy of the genetic code. SNPs occurring outside the region of a gene, or in an
20 intron within a gene, do not result in changes in any amino acid sequence of a protein but may result in altered regulation of the expression pattern. Examples include alteration in temporal expression, physiological response regulation, cell type expression regulation, intensity of expression, and stability of transcribed message.

Presented information includes that associated with genomic clones, public genes
25 and ESTs sharing sequence identity with the disclosed sequence and CuraGen Corporation's Electronic Northern bioinformatic tool.

EXAMPLES

Example A: Polynucleotide and Polypeptide Sequences, and Homology Data

30 The NOV1 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 1A.

Table 1A. NOV1 Sequence Analysis			
	SEQ ID NO: 1	750 bp	
NOV1a, CG100126-01 DNA Sequence	TCCAAACCCAAGAACTTCACTCTCTTGGAAACCCACCCAGATCCTCCCGTTCTGACA CCATGGTCAGCTCCTGTTGTGGCTCCGTGTGCTCTGACCAGGGGTGCGGCCAAGACCT CTGTGAGGAGACCTGCTGCCGCCCCAGCTGCTGTGAGACCACCTGCTGCAGGACCACC TGCTGCCGCCCCAGCTGTTGTGTATCCAGTGTGTCAGGCCCCAGTGTGTCAGTCTG TGTGCTGCCAACCACCTTGTTCGCCCCAGCTGCTGTGAGACCACCTGTTGCAGGAC CACCTGCTACCGCCCCAGCTGTTGTGTGTGCCAGTGTGTCAGGCCCGACCTGTGCCAG CCTGTGTGCTGCCAACCACCTGCTGTGCCCCAGCTGCTGTGAGACACCTGCTGCC ACCCTAGGTGCTGCATCTCCAGTGTGTCGCCCCAGCTGCTGTGTGTCCAGTGTGTG CAAGCCCCAGTGTGCCAGTCTGTGTGCTGCCAGGCCAACTGCTGCCGCCCCAGTGC AGCATCTCCAGTGTGCTGCCGCCCTCTTGCTGTGAATCCAGTGTGTCGCCGCCCTGCT ACCCTAGGTGCTGCATCTCCAGTGTGTCGCCCCAGCTGCTGTGTGTCCAGTGTGTG CAAGCCCCAGTGTGCCAGTCTGTGTGCTGCCAGGCCAACTGCTGCCGCCCCAGTGC AGCATCTCCAGTGTGCTGCCGCCCTCTTGCTGTGAATCCAGTGTGTCGCCGCCCTGCT GCTGCGTGCCTCCAGTCTGTGGCCGAGTCTCCTGCCACACCACCTTGCTATGCCCAAC CTGTGTCTATCTCCAGTGTGCCCGGCCCTTGTGCTGTGCCTCCTCTTGCTGTGAGCC CACTGCCCTGGCTTATCTCCCCCTTCACTACTGGCCACAGATGTAGACCCTTC		
	ORF Start: ATG at 61	ORF Stop: TGA at 691	
	SEQ ID NO: 2	210 aa	MW at 22433.0kD
NOV1a, CG100126-01 Protein Sequence	MVSSCCGSVCSDQCGQDLQETCCRPSCCETTCCRTTCCRPSCCVSSCCRPQCCQSV CCQPTCSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCCQPVCCQPTCCRPSCCETTCCCH PRCCISSCCRPSCCVSSCCRPQCCQSVCCQPNCCRPSCSISSCCRPSCCESSCCRPCC CVRPVCGRVSCHTTCYRPTCVISSCPRLCCASSCC		

Further analysis of the NOV1a protein yielded the following properties shown in Table 1B.

Table 1B. Protein Sequence Properties NOV1a	
Psort analysis:	0.6500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen); 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV1a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded
 several homologous proteins shown in Table 1C.

Table 1C. Geneseq Results for NOV1a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV1a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE02058	Human four disulfide core domain (FDCD)-containing protein - Homo sapiens, 230 aa. [WO200140249-A1, 07-JUN-2001]	1..209 1..204	149/218 (68%) 167/218 (76%)	e-102
ABG29368	Novel human diagnostic protein #29359 - Homo sapiens, 206 aa. [WO200175067-A2, 11-OCT-2001]	1..210 12..206	134/210 (63%) 158/210 (74%)	2e-94
ABG29368	Novel human diagnostic protein #29359 - Homo sapiens, 206 aa. [WO200175067-A2, 11-OCT-2001]	1..210 12..206	134/210 (63%) 158/210 (74%)	2e-94
ABB12277	Human hair keratin associated protein homologue, SEQ ID NO:2647 - Homo sapiens, 120 aa. [WO200157188-A2, 09-AUG-2001]	1..120 1..120	119/120 (99%) 119/120 (99%)	8e-83
AAM79986	Human protein SEQ ID NO 3632 - Homo sapiens, 301 aa. [WO200157190-A2, 09-AUG-2001]	2..210 80..301	110/227 (48%) 141/227 (61%)	2e-70

In a BLAST search of public sequence databases, the NOV1a protein was found to have homology to the proteins shown in the BLASTP data in Table 1D.

Table 1D. Public BLASTP Results for NOV1a				
Protein Accession Number	Protein/Organism/Length	NOV1a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BYQ8	KERATIN ASSOCIATED PROTEIN 4.9 - Homo sapiens (Human), 191 aa (fragment).	20..210 1..191	191/191 (100%) 191/191 (100%)	e-137
Q9BQ66	KERATIN ASSOCIATED PROTEIN 4.12 (SIMILAR TO RIKEN CDNA 1110054P19 GENE) - Homo sapiens (Human), 201 aa.	1..210 1..201	183/210 (87%) 191/210 (90%)	e-128
Q9BYR0	KERATIN ASSOCIATED PROTEIN 4.7 - Homo sapiens (Human), 210 aa.	1..210 1..210	179/215 (83%) 188/215 (87%)	e-127
Q9BYQ5	KERATIN ASSOCIATED PROTEIN 4.15 - Homo sapiens (Human), 193 aa (fragment).	13..210 1..193	171/207 (82%) 175/207 (83%)	e-114
Q9BYQ6	KERATIN ASSOCIATED PROTEIN 4.14 - Homo sapiens (Human), 195 aa.	1..210 1..195	161/210 (76%) 172/210 (81%)	e-112

Pfam analysis predicts that the NOV1a protein contains the domains shown in the Table 1E.

Table 1E. Domain Analysis of NOV1a			
Pfam Domain	NOV1a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Keratin_B2	2..108	57/177 (32%) 92/177 (52%)	6.5e-12
Keratin_B2	109..210	54/177 (31%) 82/177 (46%)	3.4e-05

Example 2.

- 5 The NOV2 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 2A.

Table 2A. NOV2 Sequence Analysis

Table 2A. NOV2 Sequence Analysis			
	SEQ ID NO: 3	3005 bp	
NOV2a, CG100146-01 DNA Sequence	CCAAACACACTAAAATAAATATGAGGTCATCAATCTTTTGTGGTCTCCTTGGCATGC ACCTATTAGACTGTTAGTATTATGTATTACTTCAAATTTTAGCAGTTATATTTTAA CTTGATTGATTTTCTCAGATATAAGTATGAGAAATGACAGAAGAAACAACACTG GAAAAGAAGCATTGCATAAGACCAGGATGTCTCTGAAATGGAGCTCAGTCTTCTGCT GATACAGCTCAGTTTTACTTTAGCTCTGGGAGTGTGGAAAGGTGCTGGTATGGCCC ACAGAATACAGCCTTTGGATGAATATGAAGACAATCCTGAAAGAACTTGTTCAGAGAG GTCATGAGGTGACTGTACTGGCATCTTCAGCTTCCATTCTTTTGTATCCCAACGACTC ATCCACTCTTAACTTGAAGTTTATCCTACATCTTTAACTAAAAATGAATTTGAGAA ATCATCATGCAATTGGTTAAGAGATGGATATATGGTGTTCAAAAAGATGCATTTTGG TTACCTTTCACAAGAACAAGAAATCCTGTGGGCAATATTATGACATAATTAGAAACTT CTGTAAGATGTAGTTTCAAATAAGAAACTTATGAAAAAACTACAAGAGTCAAGATTT GACATCGTTTTTGCAGATGCTTATTTACCCTGTGGTGAGCTGCTGGCTGAGCTATTTA ACATACCTTTTGTGTACAGTCTCCGCTTCTCTGTTGGCTACACAGTTGAGAAGAATGG TGGAGGATTTCTGTTCCCTCCTTCTATGTACCTGTGTGTATGTGAGAAATTAAGTGAT CAAATGACTTTTCATGGAGAGGGTAAAAAATATGATCTATGTGCTTTTACTTTGACTTT GGTTCGAAATATTTGACATGAAGAAGTGGGATCAGTTTTATAGTGAAGTTCTAGGAAG ACCCACTACATTATCTGAGACAATGGGGAAAAGCTGACGTATGGCTTATTTCGAAACTAC TGGGATTTTCAATTTCTCACCCTCTTACCAAATGTTGAGTTCGTTGGAGGACTCC ACTGCAAACTGCCAAACCCCTACCGAAGGAAATGGAAGAGTTTGTCCAGAGCTCTGG AGAAAATGGTGTGTGGTGTTTTCTCTGGGGTCGATGGTCAGTAACATGTGAGAAGAA AGTGCCAACATGATTGCATCAGCCCTTGCCAGATCCCAAAAAGGTTCTATGGAGAT TTGATGGCAAGAAGCCAAATACTTTAGGTTCCAATACTCGACTGTACAAAGTGGATACC CCAGAATGACCTTCTAGGTCATCCAAGACCAAAGCTTTTATAACTCATGGTGGAAACC AATGGCATCTATGAGGCAATCTACCATGGGATCCCTATGGTGGGCATTCCTTGTGTTG CGGATCAACATGATAACATTGCTCACATGAAAGCCAAGGGAGCAGCCCTCAGTGTGGA CATCAGGACCATGTCAAGTAGAGATTGCTCAATGCATTGAAGTCAGTCATTAATGAC CCTATCTATAAAGAGAAATATCATGAAATTATCAAGAATTTCATCATGATCAACCAATGA AGCCCTGGATCGAGCAGTCTTCTGGATTGAGTTTGTCTGCGCCACAAGGAGGCCAA GCACCTTCGAGTCGACGCCCAACCTCACCTGGTTCAGTACCACTCTTTGGATGTG ATAGCATTCCTGCTGGCCTGCGTGGCAACTATGATATTTATGATCACAAAATGTTGCC TGTTTGTGTTCCGAAAGTTTGCTAGAAAAGCAAAGAAGGGAAAAAATGATTAGTTATA TCTGAGATTGAAGCTGGGGAATTCCGTTTATTGAAGATTCAGGTTAACCTGAATCAA GTTAACCCAGTCTCAAATGCTCACTTATCCTTCTCTTGTGACAAATCTCTCCTCTC CTGGATTGCCAAAGAAATCAAATATTCTTCAATTAGTCAGGATGATTGACTATC AGCAGTTCATAGTACCATCTTCATAACTAAGCCACCTAGGGATCCAGCAGAAAAAAA AGGGATGAGGGGAGTCATCATACAGGAGGTGGATCATTTACCAGGATCCACACTTCCCT ACAAAGCGGTTGTAATGTTAAATAACAAAATGTTTTTTATTCCAATCTTCACATAAA ACAGGAATAATGTATACTTTCTTACTAATGTGTTCCATGGAGTTTTTCTCCAGAA GTGGCTTAGGGGAAATGAGCCCCAGTAATGCTTTGTGGCATCCAATCCTTCTACCCC GACCCCTTGAATTTCTGCCCCAGCCCTCTTAGTTCTCCTAGAAATAGGACTAAGGTT AAGTGCCCTCTTGGGATATGACTTCTTCCCTTCTCTGATACAAAAGAGCCTATT ACCAACCTCATACACACAAGAGTTCCTTCTAGTTGACAGACTCTTCTGCTCCAGCT GGACTCCCCTAGCTCTGGACTCCCACTAGATCACACAGGGGCTCCCTCATGTCAGTAA ACTTTGGATGACCTTGGGAGACAAAAATGGAATATCATTTTTGTATCTAAACAAA TAGTTTCTGATTTAACTAGGCGCCAGGAAGGTGGGCTGCACCCCTCAGTCTCTCTCC CATCATGGTTTTACATGATATCAAAGGACTCTCATAACAGTCTGATCTTATGAGTT GGGCATCTCTGTGTTTCCCTTTAGGGGCTGCTTCTCCTCAAATAGAGGAGATGGGTGCT ATGAAACCTATTCACTCTGGACTTGGGATGGCTCTTCTCCATCTTCCCAAGTCTGAGC TGGAGCCTCCATGCCCAACTCTGCTCTGCTCTTCTATTCTCTGACAGCAGCTAAGGCA TGTTCTCTGTTCTGCCCCAAATGACCTTACTAACAGTGAGAACTTGGAGGAGTCTTC GGGCTCTGGGAAATCAAAGTTTCCCGGAAACGTTTTGTTGTAAACAGTGTCCCACT CTTTGCTCCAATAAGCTCGGTTCTTAAGCCAAAAAATAAAAAA		
	ORF Start: ATG at 201	ORF Stop: TAG at 1791	
	SEQ ID NO: 4	530 aa	MW at 60923.7kD
NOV2a, CG100146-01 Protein Sequence	MSLKWTSVFLLIQLSFYSSGSCGKVLVWPTEYSLWMNMKTIKELVQRGHEVTVLAS SASILFDPNDSSSTLKLEVYPTSLTKNEFENIIMQLVKRWIYGVSKRCLLVTFHKNKKS CGQYYDIIRNFCKDVVSNKKMLKKLQESRFDIVFADAYLPCGELLAELFNIPFVYSRLR FSVGYTVEKNGGGLFPSPYVPVVMSELSDQMTFMERVKNMIYVLYPFWFEIFDMKK WDQFYSEVLGRPTLSETMGKADVWLIRNYWDFQFPHPLLPNVFVGLHCKPAKPLP KEMEEFVQSSGSENGVVFSLGSMVSNMSEESANMIASALAIQPKVLWRFDGKKPNTL GSNTRLRYKIWPQNDLLGHPKTKAFITHGGTNGIYEAIYHGI PMVGIPLFADQHDNIAH MKAKGAALSVDIRTMSRDLNALKSVINDPIYKENIMKLSRIHHDQMPKPLDRAVFW IEFVMRHKGAXHLRVAHNLTWFPQYHSLDVIAPLLACVATMIFMTKCLFCFRKFAR KAKKGKND		

Further analysis of the NOV2a protein yielded the following properties shown in Table 2B.

Table 2B. Protein Sequence Properties NOV2a	
PSort analysis:	0.8200 probability located in endoplasmic reticulum (membrane); 0.4600 probability located in plasma membrane; 0.2000 probability located in lysosome (membrane); 0.1790 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 25 and 26

A search of the NOV2a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 2C.

Table 2C. Geneseq Results for NOV2a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV2a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY78935	Human UDP-glucuronosyltransferase 2B15 amino acid sequence - Homo sapiens, 530 aa. [WO200006776-A1, 10-FEB-2000]	1..530 1..530	447/530 (84%) 485/530 (91%)	0.0
AAW47126	Uridine diphospho-glucuronosyltransferase 2B17 (UGT2B17) enzyme - Homo sapiens, 530 aa. [WO9744466-A1, 27-NOV-1997]	1..530 1..530	445/530 (83%) 480/530 (89%)	0.0
ABG05523	Novel human diagnostic protein #5514 - Homo sapiens, 533 aa. [WO200175067-A2, 11-OCT-2001]	1..530 5..533	442/530 (83%) 477/530 (89%)	0.0
ABG05523	Novel human diagnostic protein #5514 - Homo sapiens, 533 aa. [WO200175067-A2, 11-OCT-2001]	1..530 5..533	442/530 (83%) 477/530 (89%)	0.0
ABG05524	Novel human diagnostic protein #5515 - Homo sapiens, 532 aa. [WO200175067-A2, 11-OCT-2001]	1..530 4..532	446/531 (83%) 479/531 (89%)	0.0

In a BLAST search of public sequence databases, the NOV2a protein was found to have homology to the proteins shown in the BLASTP data in Table 2D.

Table 2D. Public BLASTP Results for NOV2a				
Protein Accession Number	Protein/Organism/Length	NOV2a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P54855	UDP-glucuronosyltransferase 2B15 precursor, microsomal (EC 2.4.1.17) (UDPGT) (UDPGTH-3) (HLUG4) - Homo sapiens (Human), 530 aa.	1..530 1..530	447/530 (84%) 485/530 (91%)	0.0
O75795	UDP-glucuronosyltransferase 2B17 precursor, microsomal (EC 2.4.1.17) (UDPGT) (C19-steroid specific UDP-glucuronosyltransferase) - Homo sapiens (Human), 530 aa.	1..530 1..530	445/530 (83%) 480/530 (89%)	0.0
P36537	UDP-glucuronosyltransferase 2B10 precursor, microsomal (EC 2.4.1.17) (UDPGT) - Homo sapiens (Human), 528 aa.	1..530 1..528	451/530 (85%) 484/530 (91%)	0.0
P16662	UDP-glucuronosyltransferase 2B7 precursor, microsomal (EC 2.4.1.17) (UDPGT) (3,4-catechol estrogen specific) (UDPGTH-2) - Homo sapiens (Human), 529 aa.	1..530 1..529	442/530 (83%) 477/530 (89%)	0.0
O02663	UDP-glucuronosyltransferase 2B9 precursor, microsomal (EC 2.4.1.17) (UDPGT) - Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), 529 aa.	1..530 1..529	441/530 (83%) 479/530 (90%)	0.0

PFam analysis predicts that the NOV2a protein contains the domains shown in the Table 2E.

Table 2E. Domain Analysis of NOV2a			
Pfam Domain	NOV2a Match Region	Identities/ Similarities for the Matched Region	Expect Value
UDPGT	24..528	356/507 (70%) 491/507 (97%)	0

Example 3.

The NOV3 clone was analyzed, and the nucleotide and encoded polypeptide
5 sequences are shown in Table 3A.

Table 3A. NOV3 Sequence Analysis

Table 3A. NOV3 Sequence Analysis			
	SEQ ID NO: 5	2310 bp	
NOV3a, CG100179-01 DNA Sequence	TGGCCGGCGGATAAGGGTTTCACCTTCAGGACTGGAGCTCCACAAGATTCAACCGTTAT ATCAACCTCCCCGGCCAGGTCTTACCATCACAGCGTCACAACTCCAGGTCGCCTAG GCGCTGCGCAGGAAGCGCTTGCCAGCCCCGGACTTCTGCGCGCGCTGCATGCCCATTT GATGTGCTCATTGCCACCCAGCCAATCCCGAACTCGCTCGGACGCTGACAGAAGA TGCGCGCGCTGCTTTGGGATTGGTAGCTGAGTTTGTGTCGCGCCTTTTCTGACGATG GAACAACATGGCGCGCGGAAAGTGGTAGCGATTTTCAGCAGAGACGTAGAAGGCGCCGG GACCCGGAGGAACCGGAAAAACAGAACTCAGCGAAAGAGAGCTGGCAGTAGCAGTGG CGGTGTCCCAGGAGAACGATGAGGAGAACGAAGAGCGCTGGGTTGGACCTTTACCTGT GGAGGCAACACTGGCCAAGAAGAGGAAAGTCTTAGAGTTTGAAGAGTCTATCTTGA AATCTCCCCAGTGATCCATGTATGAGCGCAGTTACATGCATAGAGATGTTATCACCC ATGTGGTATGCACCAAGACAGATTTTATTATTACTGCCAGTCATGATGGACATGTCA GTTCTGGAaaaaaaATAGAAGAGGGAATTGAATTTGTTAAACATTTTCGTAGTACCTG GGTGTATTGAGAGTATTGCAGTTAGCTCTGAGGGAGCATTGTTCTGTTCTGTGGGTG ATGATAAAGCAATGAAGGTGTTTGATGTAGTGAACCTTTGACATGATCAACATGCTGA ACTTGGGTATTTTCTGGACAGTGTGAGTGGATCTATTGCCAGGGGATGCAATTTCT TCAGTTGCTGCTTCCGAAAAGAGTACAGGAAAAATTTTCATTTATGATGGCCGAGGA ATAACCAGCCACTTCATATTTTTCAGCAAACTCCATACATCACCTCTTACTCAGATAC GCTAAACCCAGTTTACAAGCAGTAGTGTCTTCTGACAAATCTGGGATGATTGAATAC TGGACTGGGCTCCTCATGAATATAAATTCGCCAAAAATGTGAATCTGGGAATATAAA CTGACACTGATTTATATGAATTTGCCAAGTGTAAAGGCTTATCCAACAGCGTATGTT TTCACAGATGGGAAGAAAAATAGCTACTATTGGTTCTGATAGAAAAGTTAGAATTTT AGATTGTAACTGGAAAACTCATGAGAGTCTTTGATGAATCACTAAGCATGTTTACTG AACTGCAACAGATGAGGCAACAGTTACCAGACATGGAATTTGGCCGACGAATGGCTG AGAACGTGAGTTGGAGAAGGTGATGCTGTAAGATTAATTAATATAGTTTTTGATGA ACTGGACACTTCGTGCTGTATGGAACAATGCTGGGCATTAAAGTTATAAATGTAGAG CAAAACAGGTGTGCGGATTTTAGGCAAAACAGAAAAATTATGAGTGTGCAATTTGG TTTGTTCAGGGGATAGCCAAAAAGCATCGTGCTGCAACTACTATAGAAATGAAAGC TCTGAAAATCCTGTTCTTCAGAAATATTCAAGCTGACCCAACAATAGTCTGTACATCA TCAAAAAGAAATAGATTTTATATGTTTACCAACGAGAACCAAGATACGAAAAGTGC AGATTCTGATCGAGATGTTTTAATGAGAAAACTTCTAAAGAAGAAAGTCATGGCAGC ACTCAAGCTGAAGGACCTAAACGAGTTTCGGACAGTGCCATTATCCACACCAGCATG GAGACATTACACCAAACTTTTCTGTTGAGTGCCTAAGACAGTGGAAAACTTCTG TGTTACAGCAGAAAAATGGTTATTATAATGGGCATACATTTACCGTATAAATTAAGGG TTTATGATTGAGACTGGAGATCCAAACAGGTACTGGTATGGGAGGAGAAAGCATATGG GAGGAGAATTTGAAGATGAATTTCAATCAACATTACGACATGACAGGCCGTACACAC CAGCATGGCTAACCGGGGATCAAATACTAATGGATCCAGTTTTCATAACCGGTAGT CCAAACGCTTGGCTTGATAATAAGCATACAGTATTGGACGAGTGAACAAAGGAATG AAGTTGTACAGAGGATCTCCAACGTCAAAGTCAATCCCAAAACAGATAAGCCCTATG GGATGTCAGCATCATAAATATTACTGTCAAGTAAAAATAAGATTGTTTTAATGTACT GCAATAAAAAATACAATATTAACAGATTATTTTACATTAGGAAGCTT		
	ORF Start: ATG at 298	ORF Stop: TAA at 2236	
	SEQ ID NO: 6	646 aa	MW at 73574.2kD
NOV3a, CG100179-01 Protein Sequence	MAESGSDFFQRRRRRRDPEEPEKTELSERELAVAVAVSQENDEENEERWVGPLPVE TLAKKRKVLFEFERYLDNLPSASMYERSYMRDVIITHVVTCTKDFIITASHDGHVFK KKIEEGIEFVKHFRSHLGVIESIAVSSEGA LFCVSGDDKAMKVFVNVFDMINMLKL YFPGQCEWICYPGDAISSVAASEKSTGKIFIYDGRGDNQPLHIFDKLHTSPLTQIRLN PVYKAVVSSDKSGMIEYWTGPPHEYKFPKNVNWEYKTDLDLYEFAKCKAYPTSVCFSP DGKKIATIGSDRKVRI FRFVTGKLMRVFDESLSMFTELQMQRLPDMEFGRMAVE ELEKVDVRLINIVFDETHGFVLYGTMLGIKVINVETNRCVRILGQENIRVMQLALF QGIAKKHRAATTIEMKASENPVLQNIQADPTIVCTSFKNRFYMFTRKREPDTKSADS DRDVFNEKPSKEEVMAATQAEQPKRVSDSAI IHTSMGDIHTKLFVPECPTVENFCV SRNGYNGHTFHRI IKGFMIQTDPTGTGMGGESIWGGEFEDEFHSTLRHDPYTLMS ANAGSNTNGSQFFITVVP TPWLDNKHTVFGRVTKGMEVVQRI SNVKVNPKTDKPYEDV SIINITVK		

Further analysis of the NOV3a protein yielded the following properties shown in Table 3B.

Table 3B. Protein Sequence Properties NOV3a	
PSort analysis:	0.7000 probability located in nucleus; 0.4969 probability located in microbody (peroxisome); 0.3600 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV3a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 3C.

Table 3C. Geneseq Results for NOV3a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV3a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM79775	Human protein SEQ ID NO 3421 - Homo sapiens, 660 aa. [WO200157190-A2, 09-AUG-2001]	1..646 15..660	644/646 (99%) 645/646 (99%)	0.0
AAM78791	Human protein SEQ ID NO 1453 - Homo sapiens, 565 aa. [WO200157190-A2, 09-AUG-2001]	82..646 1..565	565/565 (100%) 565/565 (100%)	0.0
AAM43579	Human polypeptide SEQ ID NO 257 - Homo sapiens, 483 aa. [WO200155308-A2, 02-AUG-2001]	165..646 2..483	478/482 (99%) 478/482 (99%)	0.0
AAM43649	Human polypeptide SEQ ID NO 327 - Homo sapiens, 420 aa. [WO200155308-A2, 02-AUG-2001]	227..646 1..420	420/420 (100%) 420/420 (100%)	0.0
ABB59961	Drosophila melanogaster polypeptide SEQ ID NO 6675 - Drosophila melanogaster, 342 aa. [WO200171042-A2, 27-SEP-2001]	7..343 2..335	178/342 (52%) 241/342 (70%)	e-100

- 5 In a BLAST search of public sequence databases, the NOV3a protein was found to have homology to the proteins shown in the BLASTP data in Table 3D.

Table 3D. Public BLASTP Results for NOV3a				
Protein Accession Number	Protein/Organism/Length	NOV3a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96BP3	HYPOTHETICAL 73.6 KDA PROTEIN - Homo sapiens (Human), 646 aa.	1..646 1..646	646/646 (100%) 646/646 (100%)	0.0
Q15002	KIAA0073 PROTEIN - Homo sapiens (Human), 645 aa (fragment).	2..646 1..645	645/645 (100%) 645/645 (100%)	0.0
Q960Q8	LD39850P - Drosophila melanogaster (Fruit fly), 637 aa.	7..643 2..633	376/642 (58%) 487/642 (75%)	0.0
Q8W4D0	HYPOTHETICAL 71.0 KDA PROTEIN - Arabidopsis thaliana (Mouse-ear cress), 631 aa.	25..645 10..628	334/633 (52%) 446/633 (69%)	0.0
Q9LXM7	HYPOTHETICAL 70.7 KDA PROTEIN - Arabidopsis thaliana (Mouse-ear cress), 629 aa.	25..645 10..626	334/633 (52%) 444/633 (69%)	0.0

PFam analysis predicts that the NOV3a protein contains the domains shown in the Table 3E.

Table 3E. Domain Analysis of NOV3a			
Pfam Domain	NOV3a Match Region	Identities/ Similarities for the Matched Region	Expect Value
WD40	125..161	7/37 (19%) 29/37 (78%)	0.27
WD40	272..308	13/37 (35%) 30/37 (81%)	0.94
pro_isomerase	501..643	80/164 (49%) 115/164 (70%)	1.2e-61

Example 4.

- 5 The NOV4 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 4A.

Table 4A. NOV4 Sequence Analysis			
	SEQ ID NO: 7	1060 bp	
NOV4a, CG100212-01 DNA Sequence	<u>CAATGATTGTTCAAAGAGTGGTATTGAATCTCGACCTGGTAAAAATGGTAAATCCAGTGGCAGAGAATTTCCGAATGGAAGAAGTCTATTTACCAGATAATATTAATGAAGGACAA GTACAAGTTAGAACTCTTTATCTTTCTGTGGATCCTTACATGCGTTGTAGAATGAATGAAGACACTGGCACTGATTATATAAACACCTTGGCAGCTATCTCAAGTCGTTGATGGAGGAGGTATTGGAATTATAGAAGAAAGCAAACACACAAATTTGACTAAAGGCGATTTTGTGACTTCTTTCTATTGGCCCTGGCAAACCAAGGTTATTCTGGATGGAATAGCCTTGAAAAGGTGGACCCACAACCTTGTGGATGGACACCTTTCATATTTTCTGGAGCTATAGGTATGCCTGGTTTGACTTCCTTGATTGGGATACAGGAAAAAGGTCATATAACTGTGGATCTAATAAGACAATGGTTGTGTCAGTGGGGCCGCGAGGTGCCTGTGGATCTGTGGCTGGGCAGATTGGCCATTTCTTAGGTTGTTCCAGAGTGGTGGGAATTTGTGGAACACATGAGAAATGCATCCTCTTGACCTCAGAACTGGGCTTTGATGCTGCAATTAATTATAAAAAAGACAAATGTGGCAGAACAGCTCCGTGAATCATGCCAGCTGGAGTGGATGTTTATTTTGACAATGTTGGTGGTAAACATCAGTGATACAGTGATAAGTCAGATGAATGAGAACAGCCACATCATCCTGTGTGGTCAAATTTCTCAGTACAACAAAGATGTGCCTTATCCTCCCCCGCTATCCCTGCTATAGAGGCAATCCAGAAAGAAAGAAACATCACAAGGGAAGATTCTGGTATTAAATTATAAGACAAATTTGAGCCTGGCATTCTACAGCTGAGTCAGTGGTTTAAAGAGGAAAGCTAAAGGTAAGAGAGACGGTAATAAATGGGTGGGAAACATGGGAGCATTCAGTCCATGATGACAGGAGTAACATTGGAAAGCAGATAGTTTGCAATTCAGAAGAAATCTCTTTGTAATTGCT</u>		
	ORF Start: ATG at 3	ORF Stop: TAA at 1053	
	SEQ ID NO: 8	350 aa	MW at 38413.5kD
NOV4a, CG100212-01 Protein Sequence	MIVQRVVLNSRPGKNGNPVAENFRMEEVYLPDINNEGQVQVRTLYLSVDPYMRMRNE DTGTDYITPWQLSQVVDGGGIGIIEESKHTNLTKGDFVTSFYWPWQTKVILDGNSLEK VDPQLVDGHLSTYFLGAIGMPGLTSLIGIQEKGHITAGSNKTMVVSAGAGCGSVAGQIGHFLGCSRVVGICGTHEKCILLTSELGFDAAINYKKNVAEQRESCPAGVDVYFDNVNGGNISDVTISQMNENSHIILCGQISQYNKDVPPPLSPAIEAIQKERNITRERFLVLNYKDKFEPGILQLSQWFKGKLVKETVINGLENMGAFFQSMMTGGNIGKQIVCISEEISL		
	SEQ ID NO: 9	853 bp	
NOV4b, CG100212-02 DNA Sequence	<u>CAATGATTGTTCAAAGAGTGGTATTGAATCTCGACCTGGTAAAAATGGTAAATCCAGTGGCAGAGAATTTCCGAATGGAAGAAGTCTATTTACCAGATAATATTAATGAAGGACAA GTACAAGTTAGAACTCTTTATCTTTCTGTGGATCCTTACATGCGTTGTAGAATGAATGAAGACACTGGCACTGATTATATAAACACCTTGGCAGCTATCTCAAGTCGTTGATGGAGGAGGTATTGGAATTATAGAAGAAAGCAAACACACAAATTTGACTAAAGGCGATTTTGTGACTTCTTTCTATTGGCCCTGGCAAACCAAGGTTATTCTGGATGGAATAGCCTTGAAAAGGTAGACCCACAACCTTGTGGATGGACACCTTTCATATTTTCTGGAGCTATAGGTATGCCTGGTTTGACTTCCTTGATTGGGATACAGGAAAAAGGTCATATAACTGCTGGATCTAATAAGACAATGGTTGTGTCAGTGGGGCCGCGAGGTGCCTGTGGATCTGTGGCTGGGCAGATGAATGAGAACAGCCACATCATCTGTGTGGTCAAATTTCTCAGTACAACAAAGATGTGCCTTATCCTCCCCCGCTATCCCCGCTATAGAGGCAATCCAGAAAGAAAGAAACATCAACAGGGAAGATTTCTGGTATTAAATTATAAGACAAATTTGAGCCTGGCATTTCTACAGCTGAGTCAGTGGTTTAAAGAAGGAAGGCTAAAGATTAAAGAGACGGTAATAAATGGGTGGGAAACATGGGAGCTGCATTCCAGTCCATGATGACAGGAGGTAACTTGGAAAGCAGATAGTTTGCAATTCAGAAGAAATCTCTTTGTAATTGCT</u>		
	ORF Start: ATG at 3	ORF Stop: TAA at 846	
	SEQ ID NO: 10	281 aa	MW at 31055.3kD
NOV4b, CG100212-02 Protein Sequence	MIVQRVVLNSRPGKNGNPVAENFRMEEVYLPDINNEGQVQVRTLYLSVDPYMRMRNE DTGTDYITPWQLSQVVDGGGIGIIEESKHTNLTKGDFVTSFYWPWQTKVILDGNSLEK VDPQLVDGHLSTYFLGAIGMPGLTSLIGIQEKGHITAGSNKTMVVSAGAGCGSVAGQIMNENSHIILCGQISQYNKDVPPPLSPAIEAIQKERNITRERFLVLNYKDKFEPGILQLSQWFKGRLKIKETVINGLENMGAFFQSMMTGGNIGKQIVCISEEISL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 4B.

Table 4B. Comparison of NOV4a against NOV4b.		
Protein Sequence	NOV4a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV4b	1..183	154/186 (82%)
	1..186	158/186 (84%)

Further analysis of the NOV4a protein yielded the following properties shown in Table 4C.

Table 4C. Protein Sequence Properties NOV4a	
PSort analysis:	0.6500 probability located in cytoplasm; 0.1572 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space; 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV4a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 4D.

Table 4D. Geneseq Results for NOV4a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV4a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG64386	Human alcohol dehydrogenase 39 - Homo sapiens, 351 aa. [WO200155404-A1, 02-AUG-2001]	1..350 1..351	349/351 (99%) 350/351 (99%)	0.0
AAU33607	Pseudomonas aeruginosa cellular proliferation protein #51 - Pseudomonas aeruginosa, 345 aa. [WO200170955-A2, 27-SEP-2001]	4..347 9..344	131/347 (37%) 202/347 (57%)	4e-58
AAU33607	Pseudomonas aeruginosa cellular proliferation protein #51 - Pseudomonas aeruginosa, 345 aa. [WO200170955-A2, 27-SEP-2001]	4..347 9..344	131/347 (37%) 202/347 (57%)	4e-58
AAU34567	E. coli cellular proliferation protein #148 - Escherichia coli, 376 aa. [WO200170955-A2, 27-SEP-2001]	4..345 40..373	133/350 (38%) 203/350 (58%)	2e-56
AAU34567	E. coli cellular proliferation protein #148 - Escherichia coli, 376 aa. [WO200170955-A2, 27-SEP-2001]	4..345 40..373	133/350 (38%) 203/350 (58%)	2e-56

In a BLAST search of public sequence databases, the NOV4a protein was found to have homology to the proteins shown in the BLASTP data in Table 4E.

Table 4E. Public BLASTP Results for NOV4a				
Protein Accession Number	Protein/Organism/Length	NOV4a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8VDQ1	SIMILAR TO RIKEN CDNA B830026H24 GENE - Mus musculus (Mouse), 351 aa.	1..350 1..351	308/351 (87%) 328/351 (92%)	0.0
Q9D1W8	B830026H24RIK PROTEIN - Mus musculus (Mouse), 351 aa.	1..350 1..351	304/351 (86%) 327/351 (92%)	0.0
Q09593	HYPOTHETICAL 59.0 KDA PROTEIN M106.3 IN CHROMOSOME II - Caenorhabditis elegans, 529 aa.	6..345 27..374	161/348 (46%) 218/348 (62%)	6e-83
O34812	YFMJ PROTEIN - Bacillus subtilis, 339 aa.	4..346 6..337	146/346 (42%) 208/346 (59%)	7e-61
Q9HR85	QUINONE OXIDOREDUCTASE - Halobacterium sp. (strain NRC-1), 380 aa.	7..346 53..380	144/343 (41%) 200/343 (57%)	6e-60

PFam analysis predicts that the NOV4a protein contains the domains shown in the Table 4F.

Table 4F. Domain Analysis of NOV4a			
Pfam Domain	NOV4a Match Region	Identities/ Similarities for the Matched Region	Expect Value
adh_zinc	22..344	77/472 (16%) 216/472 (46%)	3e-16

5 Example 5.

The NOV5 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 5A.

Table 5A. NOV5 Sequence Analysis			
	SEQ ID NO: 11	978 bp	
NOV5a, CG100222-01 DNA Sequence	GATGATTGATGCTGAGATGGACCCCTGAAGAAGCACTTTGTTGGCAATCCTACTAATAG TGACTTTGAGTTGAAGACAGCTGAGCTCCACCCCTTAAAAATAGAGAGGTCTGTGCTT GAAGCTTTGTTCTCTACTGGGGACCCCTTACATGAGAGTGGCAGCCAAAGACTGAAGG AAGGTGATACAATGATGGGGCAGCAAGTGGCCAGATTGTGGAAAGTAAAAACGCAGC CTTACCAAAACAGAACTATTGTACTGGCTCCTTCAGGCTGGACAGCACACTCCATTGCT CATGGAAAAGATCTGGAAAACCTGCCAACAGAAATGGCCTGACACAGTACCCTGTCTC TGGCTCTGGGGACAGTTAGCATGACAGGCCTCACTGCCTACTTTGGCCTACTTGACAT CTGTGGTGTGAAGGGTGGAGAAACAGTGTGGTTAATGCAGCAGCTGGAGCTGTGGGC TCTGTTTTGGGGCAGATTACAAAGCTCAAGGGCTGCAAAGTGGTTGGAGCAGTAGGGT CTGACGAAAAGACTGCCTGCCTTAAAAAGCTTGGATCTGATGTCTTTAACTACAAGAC AGTAGAGTCTTTGGAAGAACTTTGAAGAAAGCCTCTCCTGATGGTTACAATTGTTAT TTTGGTAATGTAAGTGGATGGTTTTCAAACACTGTTATCTCCAGATGAAGAAATTTG GAAGAATTGCCACATGTGGAGCTATCTCTATATATAACACCAGCCCACTTCCCCCAGG CCCAATCCCAGAGATTGTTATGTATCAGGAGCTTCACATGGAAGGGTTCATTGTGACC TGCTGGCCAGGAGATGCCCAACAAAAGCTCTAAAGGACTTGCTGAAATGGGTCTCAG AGGGTCAAATCCAGTACCAGGAATATATCATTGAAGGATATGAAAACATGCCAGCTGC ATTATGGGAATGCTGAAAGGAGATAATTGGAAAAACAATAGTGAAGC		
	ORF Start: ATG at 9	ORF Stop: TAG at 969	
	SEQ ID NO: 12	320 aa	MW at 34932.1kD
NOV5a, CG100222-01 Protein Sequence	MLRWTLKKHFVGNPTNSDFELKTAELPPLKNREVLLEALFLTGDPMRVAAKRLKEGI TMMGQQVARFVESKNAALPNRTIVLAPSGWTAHSHAGKDLNLPTEWPDVPLSLAL GTVSMTGLTAYFGLLDICGVKGETVMVNAAAGAVGSVLGQITKLKGCKVVGAVGSDE KTACLKKLGSDFVNYKTVESLEETLKKASPDGYNCYFGNVSGWFSNTVISQMKKFGRI ATCGAISIYNTSPLPPGPIPEITVMYQELHMEGFIVTCWPGDAHQKALKDLKQWVSEGO IQYQEYIIIEGYENMPAAMFGLKGNLEKQ		

Further analysis of the NOV5a protein yielded the following properties shown in Table 5B.

Table 5B. Protein Sequence Properties NOV5a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.1644 probability located in microbody (peroxisome); 0.1620 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV5a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 5C.

Table 5C. Geneseq Results for NOV5a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV5a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM39169	Human polypeptide SEQ ID NO 2314 - Homo sapiens, 329 aa. [WO200153312-A1, 26-JUL-2001]	4..319 7..324	269/318 (84%) 281/318 (87%)	e-151
AAM40955	Human polypeptide SEQ ID NO 5886 - Homo sapiens, 417 aa. [WO200153312-A1, 26-JUL-2001]	4..319 60..412	267/353 (75%) 280/353 (78%)	e-145
AAU37324	Staphylococcus aureus cellular proliferation protein #1494 - Staphylococcus aureus, 334 aa. [WO200170955-A2, 27-SEP-2001]	12..320 15..328	120/319 (37%) 175/319 (54%)	5e-49
AAU37324	Staphylococcus aureus cellular proliferation protein #1494 - Staphylococcus aureus, 334 aa. [WO200170955-A2, 27-SEP-2001]	12..320 15..328	120/319 (37%) 175/319 (54%)	5e-49
AAU36657	Staphylococcus aureus cellular proliferation protein #827 - Staphylococcus aureus, 335 aa. [WO200170955-A2, 27-SEP-2001]	12..320 15..328	120/319 (37%) 174/319 (53%)	6e-49

In a BLAST search of public sequence databases, the NOV5a protein was found to have homology to the proteins shown in the BLASTP data in Table 5D.

Table 5D. Public BLASTP Results for NOV5a				
Protein Accession Number	Protein/Organism/Length	NOV5a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9H1X6	BA16L21.1.1 (NADP-DEPENDENT LEUKOTRIENE B4 12-HYDROXYDEHYDROGENASE (ISOFORM 1)) - Homo sapiens (Human), 329 aa.	4..319 7..324	269/318 (84%) 281/318 (87%)	e-151
Q14914	NADP-dependent leukotriene B4 12-hydroxydehydrogenase (EC 1.1.1.-) - Homo sapiens (Human), 311 aa (fragment).	4..305 7..310	256/304 (84%) 268/304 (87%)	e-143
O62642	15-OXOPROSTAGLANDIN 13-REDUCTASE - Sus scrofa (Pig), 329 aa.	4..319 7..324	248/318 (77%) 272/318 (84%)	e-141
Q29073	NADP-dependent leukotriene B4 12-hydroxydehydrogenase (EC 1.1.1.-) - Sus scrofa (Pig), 329 aa.	4..319 7..324	247/318 (77%) 271/318 (84%)	e-141
P97584	NADP-DEPENDENT LEUKOTRIENE B4 12-HYDROXYDEHYDROGENASE (EC 1.1.1.-) (DITHIOLETHIONE-INDUCIBLE GENE-1) - Rattus norvegicus (Rat), 329 aa.	4..319 7..324	239/318 (75%) 266/318 (83%)	e-136

PFam analysis predicts that the NOV5a protein contains the domains shown in the Table 5E.

Table 5E. Domain Analysis of NOV5a			
Pfam Domain	NOV5a Match Region	Identities/ Similarities for the Matched Region	Expect Value
adh_zinc	18..319	63/471 (13%) 206/471 (44%)	1.2e-12

Example 6.

- 5 The NOV6 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 6A.

Table 6A. NOV6 Sequence Analysis			
	SEQ ID NO: 13	696 bp	
NOV6a, CG100266-01 DNA Sequence	GTACTACATTAAATCAATGTGACAAAGTCATTTTGTCTTTATGCAAACAATCAACTTGGTATTTTCGTATAGTAACAGAAAAATGGGCTTTGCAGACGCGCGTGCCACCGGGAGCCCTCTACTATCAGCCATGGTCAACCCACCGTGTTCCTCGACATCGCTGTCCACGGTGAGCCCTTGGGTGGTATCCCCTTCGAGCTGTTTGCAGACAAGATCCCAAAGACACCAGAAAACCTTTTCATGCTCTAAGCACTGGAGAAAAAGGATTTGGTTATATAGGGTTCCTGCTTTCACAGAATTATTCCTGGGGTTATGTGTGAGCGTGGTGACTTCACAAGTCATAATGGGACTGGTGGCAAGTCCATCCAAGGGGAGAAATTTGATGACAAGAACTTCATCCTAAAGACGCAGGTCCTGGCATCTTGTCCATGGCAAATACTGGACCCAACACAAACAGTTTCCAGTGTTTCATCTGCACCTGCCAAGGCTGAGTGGTTGGATGGCAAGCATGTGGTCTTTGGCAAGGTGAAAGAAGGCATAAATATTGTGAGGCCATGGAGCGCTTTGGGTCCAGGAATGCAAGACCAGCAAGAAGATCACCATTCTGACTGTGGAAAACTCTAATACGTTTGACTGTGTGTTGATATTTATCTTAACCACCAGACTATTCTTCTGTAGCTCAGGAGAGCA		
	ORF Start: ATG at 85	ORF Stop: TAA at 625	
	SEQ ID NO: 14	180 aa	MW at 19224.9kD
NOV6a, CG100266-01 Protein Sequence	MGFADAAATGSPLLSAMVNPTVFLDIAVHGEPLGGIPFELFADKIPKTPENFHALSTGEKFGYKGS CFHRIIPGVMCQRGDFTSNHTGGKSIQGEKFDDKNFILKHAGPGILSMANTGPNTNSFQCFICTAKAEWLDGKHVVPFGKVKEGINIVEAMERFGRNGKTSKKITISDCGKL		
	SEQ ID NO: 15	545 bp	
NOV6b, CG100266-02 DNA Sequence	AAATGGGCTTTGCAGACGCGCTGCCACCGGGAGCCCTCTACTATCAGTCATGGTCAACCCACCGTGTTCCTCGACATCGCTGTCCACGGTGAGCCCTTGGGTGGTATCCCCTTCGAGCTGTTTGCAGACAAGATCCCAAAGACACCAGAAAACCTTCATGCTCTAAGCACTGAGAAAAAGGATTTGGTTATAAGGGTTCCTGCTTTCACAGAATTATTCCTGGGGTTATGTGTGAGCGTGGTGACTTCACAAGTCATAATGGCACTGGTGGCAAGTCCATCCAAGGGGAGAAATTTGATGACAAGAACTTCATCCTAAAGCACGCAGGTCCTGGCATCTTGTCCATGGCAAATACTGGACCCAAACACAAACAGTTTCCAGTGTTTCATCTGCACTGCCAAGGCTGAGTGGTTGGATGGCAAGCATGTGGTCTTTGGCAAGGTGAAAGAAGGCATAAATATTGTCGAGGCCATGGAGCGCTTTGGGTCCAGGAATGGCAAGACCAGCAAGAAGATCACCACTTCTGACTGTGGAACCTCTAA		
	ORF Start: ATG at 3	ORF Stop: TAA at 543	
	SEQ ID NO: 16	180 aa	MW at 19252.9kD
NOV6b, CG100266-02 Protein Sequence	MGFADAAATGSPLLSVMVNPTVFLDIAVHGEPLGGIPFELFADKIPKTPENFHALSTGEKFGYKGS CFHRIIPGVMCQRGDFTSNHTGGKSIQGEKFDDKNFILKHAGPGILSMANTGPNTNSFQCFICTAKAEWLDGKHVVPFGKVKEGINIVEAMERFGRNGKTSKKITISDCGKL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 6B.

Table 6B. Comparison of NOV6a against NOV6b.		
Protein Sequence	NOV6a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV6b	1..180	179/180 (99%)
	1..180	179/180 (99%)

Further analysis of the NOV6a protein yielded the following properties shown in
 5 Table 6C.

Table 6C. Protein Sequence Properties NOV6a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV6a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 6D.

Table 6D. Geneseq Results for NOV6a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV6a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU01195	Human cyclophilin A protein - Homo sapiens, 165 aa. [WO200132876-A2, 10-MAY- 2001]	17..180 1..164	141/164 (85%) 148/164 (89%)	2e-81
AAW56028	Calcineurin protein - Mammalia, 165 aa. [WO9808956-A2, 05- MAR-1998]	17..180 1..164	141/164 (85%) 148/164 (89%)	2e-81
AAR13726	Bovine cyclophilin - Bos taurus, 163 aa. [US5047512-A, 10-SEP- 1991]	18..180 1..163	140/163 (85%) 147/163 (89%)	4e-81
AAG65275	Haematopoietic stem cell proliferation agent related human protein #2 - Homo sapiens, 164 aa. [JP2001163798-A, 19-JUN-2001]	18..180 1..163	140/163 (85%) 147/163 (89%)	7e-81
AAP90431	Cyclophilin - Homo sapiens (human), 164 aa. [EP326067-A, 02-AUG-1989]	18..180 1..163	140/163 (85%) 147/163 (89%)	7e-81

In a BLAST search of public sequence databases, the NOV6a protein was found to
5 have homology to the proteins shown in the BLASTP data in Table 6E.

Table 6E. Public BLASTP Results for NOV6a				
Protein Accession Number	Protein/Organism/Length	NOV6a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC39529	SEQUENCE 26 FROM PATENT WO0132876 - Homo sapiens (Human), 165 aa.	17..180 1..164	141/164 (85%) 148/164 (89%)	5e-81
P04374	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Bos taurus (Bovine), and, 163 aa.	18..180 1..163	140/163 (85%) 147/163 (89%)	1e-80
Q9BRU4	PEPTIDYLPROLYL ISOMERASE A (CYCLOPHILIN A) - Homo sapiens (Human), 165 aa.	17..180 1..164	140/164 (85%) 147/164 (89%)	2e-80
P05092	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Homo sapiens (Human), 164 aa.	18..180 1..163	140/163 (85%) 147/163 (89%)	2e-80
Q961X3	PEPTIDYLPROLYL ISOMERASE A (CYCLOPHILIN A) - Homo sapiens (Human), 165 aa.	17..180 1..164	140/164 (85%) 147/164 (89%)	5e-80

PFam analysis predicts that the NOV6a protein contains the domains shown in the Table 6F.

Table 6F. Domain Analysis of NOV6a			
Pfam Domain	NOV6a Match Region	Identities/ Similarities for the Matched Region	Expect Value
pro_isomerase	21..180	104/179 (58%) 140/179 (78%)	3e-85

Example 7.

- 5 The NOV7 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 7A.

Table 7A. NOV7 Sequence Analysis			
	SEQ ID NO: 17	665 bp	
NOV7a, CG100427-01 DNA Sequence	GTAAATAAAATAAAGTTAAATTAATTAAAATCAAATAAAATTAAGGCTTT GCAGACACTGCCACCACCTGGGCGCCTCATACTGCCAGCCATGGTCAACCCCTACTGTGT TCTTCGACATCACCGTCGATGGCGAGCCCTTGGGCAGCGTCTCCTTCGAGCTGTTTGC AGAGAAGTTTCCAAAGACAGAAGAAAACCTTCATCTTCTGAGCACTGGAGAGAAAGGA TTTGGTTATAAAAGTTCTCTGCTTTACAGAAATTATCCAGGGTTTATGTGTCGGGGTG ATGACTTCACATGCCATAACAGCACTGGTGGCAAGTCCGTCTATGGGGAGAAGTTTGT TGACGAGAACTTCGTCCTGAAGCATAACAGATCCTGGCATCTTGTCAGGGGCAAATGCT GGACCCAGTACAAATGGTTCCAGTTTTTTACCTGCACTGCTAAGACCGAGTGGTTCG GCAAGGTGAAAGAACGCATGATTATTGTGGAGGCCATGGGGCGCTTTGCGGGGAATGC GAAGACCAGCAAGAAGATCGCTGTTGTGCTGACTGTGGGCTACTCTAATATGTTTGACTT GGGTTTATCTTAACCAGCAGACCATTCCTTCTGTAGTCCAGAGAGACCCCTCTACG CCATTTGCTCGCAGTATCCTAGAATCT		
	ORF Start: ATG at 98	ORF Stop: TAA at 566	
	SEQ ID NO: 18	156 aa	MW at 17112.3kD
NOV7a, CG100427-01 Protein Sequence	MVNPTVFFDITVDGEPLGSVSFELFAEKFPKTEENFHLSTGEKGFYKSSCFHRIIP GFMCRGDDFTCHNSTGGKSVYGEKFVDENFVLKHTDPGILSRANAGPSTNGSQFFTCT AKTEWFGKVKERMIIVEAMGRFAGNGKTSKKIAVADCGLL		

Further analysis of the NOV7a protein yielded the following properties shown in Table 7B.

Table 7B. Protein Sequence Properties NOV7a	
PSort analysis:	0.6500 probability located in plasma membrane; 0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV7a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 7C.

Table 7C. Geneseq Results for NOV7a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV7a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU01195	Human cyclophilin A protein - Homo sapiens, 165 aa. [WO200132876-A2, 10-MAY-2001]	1..156 1..164	129/164 (78%) 135/164 (81%)	5e-70
AAW56028	Calcineurin protein - Mammalia, 165 aa. [WO9808956-A2, 05-MAR-1998]	1..156 1..164	129/164 (78%) 135/164 (81%)	5e-70
AAG65275	Haematopoietic stem cell proliferation agent related human protein #2 - Homo sapiens, 164 aa. [JP2001163798-A, 19-JUN-2001]	2..156 1..163	128/163 (78%) 134/163 (81%)	2e-69
AAP90431	Cyclophilin - Homo sapiens (human), 164 aa. [EP326067-A, 02-AUG-1989]	2..156 1..163	128/163 (78%) 134/163 (81%)	2e-69
AAG03831	Human secreted protein, SEQ ID NO: 7912 - Homo sapiens, 165 aa. [EP1033401-A2, 06-SEP-2000]	1..156 1..164	128/164 (78%) 134/164 (81%)	3e-69

In a BLAST search of public sequence databases, the NOV7a protein was found to have homology to the proteins shown in the BLASTP data in Table 7D.

Table 7D. Public BLASTP Results for NOV7a				
Protein Accession Number	Protein/Organism/Length	NOV7a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC39529	SEQUENCE 26 FROM PATENT WO0132876 - Homo sapiens (Human), 165 aa.	1..156 1..164	129/164 (78%) 135/164 (81%)	1e-69
Q9BRU4	PEPTIDYLPROLYL ISOMERASE A (CYCLOPHILIN A) - Homo sapiens (Human), 165 aa.	1..156 1..164	129/164 (78%) 134/164 (81%)	3e-69
Q96IX3	PEPTIDYLPROLYL ISOMERASE A (CYCLOPHILIN A) - Homo sapiens (Human), 165 aa.	1..156 1..164	129/164 (78%) 134/164 (81%)	3e-69
P05092	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Homo sapiens (Human), 164 aa.	2..156 1..163	128/163 (78%) 134/163 (81%)	4e-69
P04374	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Bos taurus (Bovine), and, 163 aa.	2..154 1..161	127/161 (78%) 133/161 (81%)	1e-68

PFam analysis predicts that the NOV7a protein contains the domains shown in the Table 7E.

Table 7E. Domain Analysis of NOV7a			
Pfam Domain	NOV7a Match Region	Identities/ Similarities for the Matched Region	Expect Value
pro_isomerase	5..154	94/177 (53%) 126/177 (71%)	7.9e-61

Example 8.

- 5 The NOV8 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 8A.

Table 8A. NOV8 Sequence Analysis			
	SEQ ID NO: 19	2646 bp	
NOV8a, CG100456-01 DNA Sequence	ATGTGGCTGCAGCAGCGGCTCAAGGGGCTGCCGGGACTGCTGTGCGAGCAGCTGGGCCC GCCGCCTCTCTGCTGCTTGGCCTCTGCTGCTGCTTCTGTGGTTTGGGGGGTCCGG CGCGCGGCGGGCGGCGGGCGGCTGCACCTGCTGCCCTGGTCCCGCGGTGAGCCGGGC GCCGCCGAGCCGTCTGCTGCTGGAGGCGGCCACCCGCGCTGGCGCGCTGCGGG AGCGCGGTGAGGTGGTACCCTGGGTCTTGGAGTGCCGGCCCTGGTGGCCAAACGGCTT CCTGGCCCTGGACGTGGCTGCCAATCGGCTGTGGGTGACTCCCGGGAGCGGGAGCCC GCCGTGGCGCCGAGACTTTGTGCCCTTCGTGCAGCTGCGCCCGCTGAGCGCGCTGGGTG AAGCTGGAGAGGCGGTGCTGCTGCTGCGGGAGGGGCTTCTGCGCCGCTGCGTGTGCCT GCAGCTGGGGTCCCCAGGTCTTGGCCCCGTGGCCGCGGCCCGGGCCCGCTCCGCT TCTGGCCTTGCCGCGGGGTCCGGCCGCGACTGCGTGCTGCTGCAAGAGGACTTCTGG CGCACAGGGGCGGACCCACGTCTACCTGCAGCGCATCCAGCTCAACAACCCACGGA GCGCGTGGCCGCGCTGCAGACTGTGGGGCCACTGCCGGGCCACGCCCAATGCCTTC ACCAGTACCCTGGAGAAGGTCCGAGACCATCAGTTCTCTCTACTCAGGCCGGTCCC CGCCTACGCCCACTGGGTGGTGCACTGGTGGTGGTGGCGCCCAAGAGCTGTGAA CCGCTTCCAAGTGGCTCCCAAGACGAGCTGGATGAGACGGTGTGTGGGTGGTGCA GTCTCTGGCCCCATTAAACCCAGGTGCTCAAAGCAAAGCAGCAAGGAGCTCAAGG CGCTGCAGGACTTGGCAGGAAGGAAATGCTGGAGCTCTTGGACATGCCAGCGGGGA GCTGCTTCAAGACACCCAGCTCTCTGGGCTCAGCTCTTCAAGCCAGGTGTGGAATG AAGAAGATCACTGACACCCACAGCCGTCTGGCCTCACCGTGAACCTGACGCTCTATT ACATGCTCTCTGCTCGCCAGCCCACTGCTCAGCCCTCCCTGAGCCACAGGGAGCG AGACCAGATGGAGTCGACGCTCAACTATGAAGATCACTGCTTCAAGCGGCACGCCACC ATGCACGCGGAGAACCTGTGGCCGGGGCGGCTGTCTCCGTCCAGCAGATCTGCAGC TCTCTGACCTGTGGAGGCTGACCTCCAGAAGCGTGGCTGCAAGGGGCTGGTGAAGGT GGGTGCCCCAGGCATCTTGCAGGGGATGGTGTCTCAGCTTTGGGGGGCTGCAGTTTACA GAGAACCACTCCAGTTCAGTTCAGGCCGACCCGACGTGCTGCACAACAGCTATGCAATG ATGGCATCCGCTACAAGAACGACCATATCAACCTGGCCGTGCTGGCGGATGCCGAGGG CAAGCCCTACCTACAGTGTCCGTGGAGTCCCGTGGCCAGCCTGTCAAGATCTATGCC TGCAAGGCAGGCTGCTGGACGAGCCAGTGGAGCTGACCTCGGCCCAACCGGCCACA CCTTCTCGGTGATGGTGACACAGCCCATCAGCCACTGCTCTACATCTCCACCGACCT CACACACCTGCAGGACCTGCGGCACACGCTGCACCTCAAGGCCATCCTGGCCCATGAT GAGCACATGGCCCCAGCAGGACCCCGGCTGCCCTTCTCTTCTGGTTGAGCTGGCCCT CCCTAATACCCCTTCTCCACCTTCTCTTCAAGCTCATCAACAGGATCTGTGG GCCTGGAGCCAAGCCCTTTCAGGAGTAAGGAAGATCCAGTGTCTGAGTGAAGTAA CAGTCTCTGCTTTAGCCACCATTTGCACAAGACACCCAGCACTGAAAGTCCCGCTGCC AGGAGCAAGGGATCCTTTGGAAGCACCCGCCCTTTGTGCTTGTGGGGGAAACCGGT GACGCAAGAGTGAAGTGGATACACAGAGTTTGCAATGGAAGGAATGAGTGTACGT GGGGAGGGAAGGGGCCAGTGGACCTTTTGAAGCTTTCCACTCAATAAATGAACCTG TATGGCAAATACTTGAATGGAACCTCACTCCTTCCACTTTCCCTTTCTTCTGTGCC AGGAATAGATCATCTTTGAAAGACTCTTGTCTAGGAAAGTGTGTGCTTTCTCTCT AATTTAAGTGTCTTCTTCTAATGAAGTTTAAATTTATTTTGTGTGAGATTTTGCTAG ATGCTCTTTGCATCCCTGTAGATGGTGAAGTGTGGCGGTGATGTCGCTCGCGGT CGGAGGCCCAAGGTCCCGAGGCTGGGCGGGGCCCCAGGGTGGCTGTGCTGTG CTGTAGGAGGGTGGGGTTGTGCTGTATCCTCGGGTTTGCACGCCCTTTTATAGGAG CCTGTGGACATCTGTGGTTTGTACTTTGGGGCTTCAGGGGAGGTGTTTAACTTTCTA GTGATTGATGATTGTGAGGTTTGAATAACCAAGCTTTTGTGTCTGTTTTAAATA AATATCTTTCAAACCTTCAAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 1	ORF Stop: TGA at 1903	
	SEQ ID NO: 20	634 aa	MW at 69142.3kD
NOV8a, CG100456-01 Protein Sequence	MWLQRLKGLPGLLSSSWARRLLCLLGLLLLLLWFGGSGARRAAGLHLLPWSRGEPP AAEPSACLEAATRAWRLRERGEVPLPGVPVLPALVANGFLALDVAANRLWVTPGEREP AVAPDFVFPVQLRPLSALAEAGEAVLLREGLRRVRCLQLGSPGPGVAAAGPGPASV SGLAAGSGRDCVLLQEDFLAHRGRPHVYLQRIQLNNPTERVAALQTVGPTAGPAPNAF TSTLEKVGHDQFLYSGRSPPTPTGLVHLVVVAAKLVNRLQVAPKTQLDETVLWVVH VSGPINPQVLKSKAKELKALQDLARKEMLELLDMPAAELLDHQLLWALQFSPGVEM KKITDTHTPSGLTVNLTLTYMLSCSPAPLLSPSLSHRERDQMETLNYEDHCFSGHAT MHAENLWPGRLSSVQQILQLSDLWRLTLQKRGCKGLVKVGAPGILQGMVLSFGLQFT ENHLQFQADPDVLHNSYALHGIRYKNDHINLAVLADAEKPYLHVSVESRGPVKIYA CKAGCLDEPVELTSAPTGHTFSVMVTQPIITPLLYISTDLTHLQDLRHTLHLKAILAHD EHMAQQDPGLPFLFVFSVASLITLFLFLFKLIYNEYCGPGAKPLFRSKEDPSV		

Further analysis of the NOV8a protein yielded the following properties shown in Table 8B.

Table 8B. Protein Sequence Properties NOV8a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.3000 probability located in lysosome (membrane); 0.2800 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 41 and 42

A search of the NOV8a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 8C.

Table 8C. Geneseq Results for NOV8a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV8a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB43028	Human ORFX ORF2792 polypeptide sequence SEQ ID NO:5584 - Homo sapiens, 454 aa. [WO200058473-A2, 05-OCT-2000]	182..634 2..454	433/453 (95%) 433/453 (95%)	0.0
ABB63158	Drosophila melanogaster polypeptide SEQ ID NO 16266 - Drosophila melanogaster, 606 aa. [WO200171042-A2, 27-SEP-2001]	111..627 96..602	155/557 (27%) 259/557 (45%)	5e-55
AAB27712	Human secreted protein #31 - Homo sapiens, 108 aa. [WO200055201-A1, 21-SEP-2000]	456..524 30..98	69/69 (100%) 69/69 (100%)	2e-34
AAB27756	Protein fragment encoded by gene 31 - Homo sapiens, 65 aa. [WO200055201-A1, 21-SEP-2000]	524..585 4..65	61/62 (98%) 62/62 (99%)	5e-29
AAB27755	Sequence homologous to protein fragment encoded by gene 31 - Homo sapiens, 65 aa. [WO200055201-A1, 21-SEP-2000]	523..585 2..65	27/66 (40%) 37/66 (55%)	3e-05

In a BLAST search of public sequence databases, the NOV8a protein was found to have homology to the proteins shown in the BLASTP data in Table 8D.

Table 8D. Public BLASTP Results for NOV8a				
Protein Accession Number	Protein/Organism/Length	NOV8a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q91X21	RIKEN CDNA 2510039O18 GENE - Mus musculus (Mouse), 634 aa.	1..634 1..634	582/634 (91%) 607/634 (94%)	0.0
Q9CY11	2510039O18RIK PROTEIN - Mus musculus (Mouse), 634 aa.	1..634 1..634	581/634 (91%) 607/634 (95%)	0.0
Q9BSY1	SIMILAR TO RIKEN CDNA 2510039O18 GENE - Homo sapiens (Human), 238 aa (fragment).	400..634 4..238	235/235 (100%) 235/235 (100%)	e-138
Q9VQ60	CG7289 PROTEIN - Drosophila melanogaster (Fruit fly), 606 aa.	111..627 96..602	155/557 (27%) 259/557 (45%)	1e-54
Q95RT9	LD12115P - Drosophila melanogaster (Fruit fly), 637 aa.	111..627 96..633	155/566 (27%) 269/566 (47%)	2e-54

PFam analysis predicts that the NOV8a protein contains the domains shown in the Table 8E.

Table 8E. Domain Analysis of NOV8a			
Pfam Domain	NOV8a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 9.

- 5 The NOV9 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 9A.

Table 9A. NOV9 Sequence Analysis			
	SEQ ID NO: 21	905 bp	
NOV9a, CG100466-01 DNA Sequence	CCTGCCTCCTCTTTCTTCAACATGACAGATGCCGCTGTGTCTTCGCCAAGGACTT CCTGGCAGGTGGAGTGGCCGAAGCCATCTCCAAGACAGCGGTAGCGCCATCGAGCAG GTCAAGCTGCTGCTGCAGGTGCAGCATGCCAGCAAGCAGATCACCTCAGCTAAGCAAT ACAAGGGCATTATAGACTGCGTGGTCCGTATTCCCAAGAAGCAGGGAGTCTGTCTCTG GCGCGGTAACCTGGCCAATGCCATCAGATACTTCCCCACCAGGCTTTAACTTCGCC TTCAAAGATAAATAACAAGCAGATCTTCTGGGTGGTGTGGACAAAAGAACTCAGTTTG GGCGCTACTTTGCAGGGAATCTGGCATCAGGAGGTGCCGCTGGGGCCACATACTTTGTG TTTTGTGTACCCTCTTGATTTTGCCCGTACCTGTCTAGCAGCTAATGTGGGTAAAGCT GAAGCTGAAAGGGAATCCGAGGCCCTCTGTGACTGCCTGGTTAAGATCTACAAATCTG ATGGGATTAAGGGCCTGTACCAAGGCTTTAACGTGTCTATGCAGGGTATTATCCGAGC TGCCTACTTCGGTATCTATGACACCGCAAAGGGAATGCTTCCGGATCCCAAGGACACT CACATCGTCATCAGCTGGATGACCACACAGACTGTCACTGCCTTTTCTGGGTGACTT CCTATTCA11TGACATCGTTTCGCGTGATGATTCACTCAGGCGCAAAGTAACTGACAT CATGTACACAGGCACACTTGACTGCTGGAGGAAGATTGCTGGTGATGAAGGAGGCAA GCTTTTTTCAAGGGTTCATGTTCCAGTGTTCTCAGAGGCATGGGTGGTGCTTTTGTGC TTGTCTTGT11TGATGAAATCAAGAAGTACAGGTAA		
	ORF Start: ATG at 24	ORF Stop: TAA at 903	
	SEQ ID NO: 22	293 aa	MW at 32215.1kD
NOV9a, CG100466-01 Protein Sequence	MTDAAVSFAKDFLAGGVAEAIKTA VAPIEQVKLLQVQHASKQITS AKQYKGI IDC V VRIPKKQGVLSWRGNLANAIRYFPTQAFNFAFKDKYKQIFLGGVDKRTQFGRYFAGNL ASGGAAGATYLCFVYPLDFARTCLAA NVGKA EAEREFRGLCDCLVKIYKSDGIKGLYQ GFNVSMQGIIRAA YFGIYDTAKGMLPDPKDT H I V I S W M T T Q T V A F S G L T S Y S F D I V R VMIQSGRKVTDIMYTGTLDCWRKIAGDEGGKAFFKGSWSSVLRMGGAFLVLFDEIK KYR		

Further analysis of the NOV9a protein yielded the following properties shown in Table 9B.

Table 9B. Protein Sequence Properties NOV9a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.2400 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV9a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 9C.

Table 9C. Geneseq Results for NOV9a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV9a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU10379	Human adenine nucleotide translocator 2 (ANT2) - Homo sapiens, 298 aa. [WO200185944-A2, 15-NOV-2001]	1..292 1..297	262/297 (88%) 275/297 (92%)	e-145
AAU01199	Human adenine nucleotide translocator-2 (ANT-2) protein - Homo sapiens, 298 aa. [WO200132876-A2, 10-MAY-2001]	1..292 1..297	262/297 (88%) 275/297 (92%)	e-145
AAU71032	Human adenine nucleotide translocator ANT2 - Homo sapiens, 298 aa. [WO200026370-A2, 11-MAY-2000]	1..292 1..297	262/297 (88%) 275/297 (92%)	e-145
AAU10380	Human adenine nucleotide translocator 3 (ANT3) - Homo sapiens, 298 aa. [WO200185944-A2, 15-NOV-2001]	1..291 1..296	242/296 (81%) 267/296 (89%)	e-137
AAU01200	Human adenine nucleotide translocator-3 (ANT-3) protein - Homo sapiens, 298 aa. [WO200132876-A2, 10-MAY-2001]	1..291 1..296	242/296 (81%) 267/296 (89%)	e-137

In a BLAST search of public sequence databases, the NOV9a protein was found to have homology to the proteins shown in the BLASTP data in Table 9D.

Table 9D. Public BLASTP Results for NOV9a				
Protein Accession Number	Protein/Organism/Length	NOV9a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P05141	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Homo sapiens (Human), 298 aa.	1..292 1..297	264/297 (88%) 276/297 (92%)	e-146
A29132	ADP,ATP carrier protein T2 - human, 298 aa.	1..292 1..297	262/297 (88%) 275/297 (92%)	e-145
Q09073	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Rattus norvegicus (Rat), 298 aa.	1..292 1..297	261/297 (87%) 275/297 (91%)	e-145
P51881	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Mus musculus (Mouse), 298 aa.	1..292 1..297	260/297 (87%) 274/297 (91%)	e-144
BAB84673	ADENINE NUCLEOTIDE TRANSLOCATOR 2 - Bos taurus (Bovine), 298 aa.	1..292 1..297	259/297 (87%) 274/297 (92%)	e-144

PFam analysis predicts that the NOV9a protein contains the domains shown in the Table 9E.

Table 9E. Domain Analysis of NOV9a			
Pfam Domain	NOV9a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	7..104	35/125 (28%) 86/125 (69%)	6.8e-28
mito_carr	111..205	37/125 (30%) 80/125 (64%)	1.2e-18
mito_carr	206..293	22/125 (18%) 61/125 (49%)	7.6e-07

Example 10.

- 5 The NOV10 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 10A.

Table 10A. NOV10 Sequence Analysis			
	SEQ ID NO: 23	858 bp	
NOV10a, CG100609-01 DNA Sequence	ACATAGGCCCCAGGTCTGTGGCTGGCCAGCTGCAGCTGTGGGGCCTGGCATGTGTCTCA ACATGGCCCTGGAGCTCTACATGGACCTGCTGTGACGACCCTGCCGTGCCGTCTACAT CTTCTCGAAGAAGCATGACATCCAGTTCACITTCAGTTTGTGGATCTGCTGAAAGGT CACCACCACAGCAAAGAATACATTGACATCAACCCCTCAGGAAGCTTCCCAGCCCTCA AAGATGGGAAATTTATCTTAAGTGAAAGCGCGCCATCCTTTACTACCTGTGCCGCAA GTACAGCGCACCATCGCACTGGTGGCCGCCAGACCTGCACGCACGTGCCCGTGTGGAT GAGTTCTGTGGCTTGGCAACACACGGCCTTTCAGCTGCCCATGAAGAAGATAGTCTGGC TCAAGTTGCTGATCCCAAGATAACAGGGGAGGAAGTTTCAGCTGAGAAGATGGAGCA TGCACTGGAAGAGGTGAAGAACAGCCTGCAGCTCTTTGAGGAGTATTTCTGCAGGAT AAGATGTTTCATCACCGGGAACCAAATCTCACTGGCTGACCTGGTGGCCGTGGTGAGGA TGATGCAGCCCATGGCAGCCAATATAATGTCTTCTCAACAGCTCCAAGCTAGCTGA GTGGCGTATGCAGGTGGAGCTGAATATTGGCTCTGGCCTCTTTAGGGAGGCCCATGAT CGACTAATGCAGTTGGCCGACTGGGACTTTTCAACATTGGATTCAATGGTCAAGGAGA ATATTTCTGAGTTGTCTGAAGAAGAGCAGGTGACCCTAAGCCGAGCCTGTCCCGCAGGG CCTGGCTGGCTTAGCAATCTGAGCCACCTTCCTTAAAGGAAATGTT		
	ORF Start: ATG at 49	ORF Stop: TGA at 784	
	SEQ ID NO: 24	245 aa	MW at 28490.9kD
NOV10a, CG100609-01 Protein Sequence	MCLNMALELYMDLLSAPCRAVYIFSKKHDIQNFQFVDLLKGHHHSKEYIDINPLRK PSLKDGGKFI LSESAAILYYLCRKYSAPSHWCPPDLHARARVDEFVAVQHTAFQLPMKH IVWLKLLIPKITGEEVSAEKMEHAVEEVKNSLQLFEEYFLQDKMFI TGNQISLADLVA VVENMQPMANYNVFLNSSKLAEWRMQVELNIGSGLFREAHDRMLQLADWDFSTLDSN VKENISELLKKS		

Further analysis of the NOV10a protein yielded the following properties shown in Table 10B.

Table 10B. Protein Sequence Properties NOV10a	
PSort analysis:	0.4826 probability located in microbody (peroxisome); 0.4708 probability located in mitochondrial matrix space; 0.1732 probability located in mitochondrial inner membrane; 0.1732 probability located in mitochondrial intermembrane space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV10a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 10C.

Table 10C. Geneseq Results for NOV10a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV10a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB85772	Human drug metabolizing enzyme (ID No. 7472055CD1) - Homo sapiens, 241 aa. [WO200159127-A2, 16-AUG-2001]	5..245 1..241	237/241 (98%) 238/241 (98%)	e-136
AAV07034	Breast cancer associated antigen precursor sequence - Homo sapiens, 240 aa. [WO9904265-A2, 28-JAN-1999]	5..237 1..231	116/233 (49%) 156/233 (66%)	9e-63
AAB54309	Human pancreatic cancer antigen protein sequence SEQ ID NO:761 - Homo sapiens, 255 aa. [WO200055320-A1, 21-SEP-2000]	5..237 15..246	111/233 (47%) 152/233 (64%)	5e-59
AAU30356	Novel human secreted protein #847 - Homo sapiens, 420 aa. [WO200179449-A2, 25-OCT-2001]	5..223 107..331	93/226 (41%) 132/226 (58%)	1e-38
AAM03015	Peptide #1697 encoded by probe for measuring breast gene expression - Homo sapiens, 64 aa. [WO200157270-A2, 09-AUG-2001]	181..244 1..64	62/64 (96%) 63/64 (97%)	5e-29

In a BLAST search of public sequence databases, the NOV10a protein was found to have homology to the proteins shown in the BLASTP data in Table 10D.

Table 10D. Public BLASTP Results for NOV10a				
Protein Accession Number	Protein/Organism/Length	NOV10a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9D4P7	4930583C14RIK PROTEIN - Mus musculus (Mouse), 240 aa.	5..243 1..238	181/239 (75%) 210/239 (87%)	e-103
Q969K8	GLUTATHIONE S-TRANSFERASE TT1 (GLUTATHIONE S-TRANSFERASE THETA 1) (GLUTATHIONE TRANSFERASE T1-1) (EC 2.5.1.18) - Homo sapiens (Human), 240 aa.	5..237 1..231	117/233 (50%) 157/233 (67%)	3e-63
Q96IY3	GLUTATHIONE S-TRANSFERASE THETA 1 - Homo sapiens (Human), 240 aa.	5..237 1..231	117/233 (50%) 156/233 (66%)	1e-62
S44358	glutathione transferase (EC 2.5.1.18) theta-1 [validated] - human, 240 aa.	5..237 1..231	116/233 (49%) 156/233 (66%)	2e-62
P30711	Glutathione S-transferase theta 1 (EC 2.5.1.18) (GST class-theta) - Homo sapiens (Human), 239 aa.	7..237 2..230	115/231 (49%) 155/231 (66%)	9e-62

PFam analysis predicts that the NOV10a protein contains the domains shown in the Table 10E.

Table 10E. Domain Analysis of NOV10a			
Pfam Domain	NOV10a Match Region	Identities/ Similarities for the Matched Region	Expect Value
GST_N	4..80	26/90 (29%) 57/90 (63%)	2.7e-12
GST_C	90..199	23/115 (20%) 75/115 (65%)	5.2e-09

Example 11.

- 5 The NOV11 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 11A.

**MISSING AT THE TIME OF
PUBLICATION**

	ORF Start: ATG at 127	ORF Stop: TAG at 721	
	SEQ ID NO: 28	198 aa	MW at 21674.4kD
NOV11b, CG100631-02 Protein Sequence	MAELDPFGAPAGAPGGPALGNGVAGAGEEDPAAAFLAQQESEIAGIENDEAFAILDGG APGPQPHGEPGGPDVADGVMNGEYYQESNGPTDSYAAISQVDRLOSEPESIRKWREE QMERLEALDANSRKQEAWEKKAIKELEEWYARQDEQLQKTKANNSTNINHPCYSLEQ AAEEALHSHWINYVEFFWTKPFCL		
	SEQ ID NO: 29	735 bp	
NOV11c, CG100631-04 DNA Sequence	CCCAGTCGGCACCACAGCGGTGGCTGCCGGCGTGGTGTGGTGGGTTGGTTGGTTTT TGCTCACCCTTGGTGTCCGTGCCGTTTCAGTTGCCCGCCATGGCTGAGCTGGATCCGT TCGGCGCCCTGCCGCGCCCTGCCGCTCCGCGTGGGGAACGAGTGGCCGCGCGC CGGCGAAGAAGACCCGCGCGGCTTCTTGGCGCAGCAAGAGAGCGAGATTGCGGGC ATCGAGAACGACGAGGCCTTCGCCAATTCTCGGAAGCAAGAAGCAGAGTGGAAAGAAA AGGCAATAAAGGAGCTAGAAGAATGGTATGCAAGACAGGACGAGCAGCTACAGAAAAC AAAAGCAAACAACAGGGCAGCAGAAGAAGCCTTTGTAATGACATTGACGAGTCGTCC CCAGGCACTGAGTGGGAACGGGTGGCCCGCTGTGTGACTTTAACCCCAAGTCTAGCA AGCAGGCCAAAGATGTCTCCCGCATGCGCTCAGTCCTCATCTCCCTCAAGCAGGCCCC GCTGGTGCACTGAAGAGCCACCCTGTGGAAACACTACATCTGCAATATCTTAATCCTA CTCAGTGAAGCTCTTACAGTCATTGGATTAAATTATGTTGAGTTCTTTTGGACCAAAAC CTTTTGTCTTTAGAGTTGTTTCATTGTTTGTGATTGCATGTTTCCTTCCTCAACTGT GTTCTGCCTGGCATTGAGAGAGGGAGAGGAGGAAGA		
	ORF Start: ATG at 98	ORF Stop: TGA at 533	
	SEQ ID NO: 30	145 aa	MW at 15769.2kD
NOV11c, CG100631-04 Protein Sequence	MAELDPFGAPAGAPGGPALGNGVAGAGEEDPAAAFLAQQESEIAGIENDEAFANSRKQ EAEWKEKAIKELEEWYARQDEQLQKTKANNRAAEAFVNDIDESSPGTEWVERVLCD FNPKSSKQAKDVSRMRSVLISLKQAPLVH		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 11B.

Table 11B. Comparison of NOV11a against NOV11b and NOV11c.		
Protein Sequence	NOV11a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV11b	1..161 1..161	125/161 (77%) 125/161 (77%)
NOV11c	126..236 53..145	93/111 (83%) 93/111 (83%)

Further analysis of the NOV11a protein yielded the following properties shown in Table 11C.

Table 11C. Protein Sequence Properties NOV11a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV11a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 11D.

Table 11D. Geneseq Results for NOV11a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV11a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB71527	Drosophila melanogaster polypeptide SEQ ID NO 41373 - Drosophila melanogaster, 219 aa. [WO200171042-A2, 27-SEP-2001]	20..234 4..213	77/231 (33%) 118/231 (50%)	2e-27
AAB57165	Human prostate cancer antigen protein sequence SEQ ID NO:1743 - Homo sapiens, 116 aa. [WO200055174-A1, 21-SEP-2000]	10..79 7..77	39/71 (54%) 42/71 (58%)	9e-11
AAV87249	Human signal peptide containing protein HSPP-26 SEQ ID NO:26 - Homo sapiens, 82 aa. [WO200000610-A2, 06-JAN-2000]	161..180 18..37	19/20 (95%) 20/20 (100%)	0.001
ABB62631	Drosophila melanogaster polypeptide SEQ ID NO 14685 - Drosophila melanogaster, 305 aa. [WO200171042-A2, 27-SEP-2001]	89..171 199..282	26/86 (30%) 47/86 (54%)	0.007
AAV71045	Streptococcus pyogenes strain AP49 partial GRAB protein - Streptococcus pyogenes, 271 aa. [WO200026240-A2, 11-MAY-2000]	73..219 98..253	46/160 (28%) 69/160 (42%)	0.025

In a BLAST search of public sequence databases, the NOV11a protein was found to have homology to the proteins shown in the BLASTP data in Table 11E.

Table 11E. Public BLASTP Results for NOV11a				
Protein Accession Number	Protein/Organism/Length	NOV11a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P09496	Clathrin light chain A (Lca) - Homo sapiens (Human), 248 aa.	1..236 1..248	236/248 (95%) 236/248 (95%)	e-133
LRRTA1	clathrin light chain A1 - rat, 286 aa (fragment).	1..236 39..286	231/248 (93%) 234/248 (94%)	e-130
P08081	Clathrin light chain A (Lca) - Rattus norvegicus (Rat), 248 aa.	1..236 1..248	231/248 (93%) 234/248 (94%)	e-130
O08585	Clathrin light chain A (Lca) - Mus musculus (Mouse), 235 aa.	1..236 1..235	220/236 (93%) 223/236 (94%)	e-123
P04973	Clathrin light chain A (Lca) - Bos taurus (Bovine), 243 aa.	1..236 1..243	223/248 (89%) 226/248 (90%)	e-121

PFam analysis predicts that the NOV11a protein contains the domains shown in the Table 11F.

Table 11F. Domain Analysis of NOV11a			
Pfam Domain	NOV11a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Clathrin_lg_ch	3..235	154/257 (60%) 226/257 (88%)	5.2e-142

Example 12.

- 5 The NOV12 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 12A.

Table 12A. NOV12 Sequence Analysis			
	SEQ ID NO: 31	1849 bp	
NOV12a, CG100710-01 DNA Sequence	CATGTCGTGGCTCTTCGGCGTTAACAAGGGCCCCAAGGGTGAAGGCGCGGGGCCGCGG CCGCCTTTGCCGCCCGCGCAGCCCGGGGCCGAGGGCGCGGGGACCGCGTTTGGGAG ACCGGGCCGGCGCCCAAGGACAAATGGAGCAACTTCGACCCACCGGCTGGAGCGCGC CGCCAAGGCGCGCGCGAGCTGGAGCACTCGCGTCAGACACAGGAGCGGCTGTGAGGC AGTGCCAGCCCTGAGCAAGTGCCAGCTGGTGAAGTGTGTCTCTGTCGAGAGTATGAGG CCGCCGTGGAGCAGCTCAAGAGCGAGCAGATCCGGGCGCAGGCTGAGGAGAGGAGGAA GACCTTGAGCGAGGAGACCCGGCAGCACCAGGCCAGGGGCCAGTATCAAGACAAGCTG GCCGGCAGCGCTACGAGGACCAACTGAAGCAGCAGCAACTTCTCAATGAGGAGAATT TACGGAAGCAGGAGGAGTCCGTGCAGAAGCAGGAAGCCATGCGGCGAGCCACCGTGGA GCGGGAGATGGAGCTGCGGCACAAGAATGAGATGCTGCGAGTGGAGACCGAGGCCCGG GCGCGCGCCAAGGCCGAGCGGGAGAATGCAGACATCATCCGCGAGCAGATCCGCTGA AGGCGTCCGAGCACCCTGAGACCGTCTTGGAGTCCATCAGGACGGCTGGCACCTTGT TGGGGAAGGATTCCGTGCCTTTGTGACAGACCGGGACAAAGTGACAGCCACGGTAGCT GGGCTGACGCTGCTGGCTGTGCGGGTCTACTCAGCCAAGAATGCGACAGCCGTCACTG GCCGCTTCATCGAGGCTCGGCTGGGGAAGCGCTCCTTAGTGAGGAGACGTCCCGCAT CACGGTGCTGGAGGCGCTGCGGCACCCCATCCAGGTAAGCCGGCGGCTCCTCAGTCGA CCCCAGGACGTGCTGGAGGGTGTGTGCTTAGTCCCAGCCTGGAAGCACGGGTGCGCG ACATCGCCATAGCAACCAAGGAACCAAGAAGAACCAGGGGCTGTACAGGCACATCCT GCTGTATGGGCCACCAGGCACCGGGAAGACGCTGTTTGCCAAGAAACTCGCCCTGCAC TCAGGCATGGACTACGCCATCATGACAGCGGGGACGTGGCCCCATGGGGCGGGAAG GCGTGACCGCCATGCACAAGCTCTTTGACTGGGCCAATACCAGCCGGCGCGGGCTCCT GCTCTTCATGAGTGAAGCAGACGCTTCTTTCGGAAGCGAGCCACTGAGGAGATAAGC AAGGACCTCAGAGCCCACTGAACGCTTCTCTGTACCACATGGGCCAACACAGCAACA AGTTCATGCTGGTCTGGCCAGCAATCTGCTGAGCAGTTTCCACTGTGCCATCAACAG CCGCAATTGACGTGATGGTCCACTTCGACCTGCGCAGCAGGAGGAGCGGGAGCGCCTG GTGAGACTGCATTTTGACAACCTGTGTTCTTAAGCCGGCCACAGAAGGAACCGCGCC TGAAGCTGGCCAGTTTGTACTACGGGAGGAAGTGCTCGGAGGTGCTCGGCTGACGGA GGGCATGTGCGGCCCGGAGATCGCTCAGCTGGCCGTGCTTGGCAGGCCACCGCGTAT GCCTCCGAGGACGGGGTCTGACCGAGGCCATGATGGACACCCGCGTGAAGATGCTG TCCAGCAGCACCAGCAGATGATGCGCTGGCTGAAGAGGGGGAGGCTGGGCCCGAGGA CGAGCAACCCTCATCTGAGTCCATGGGGAGACCACCTCAGCGAGCCTGGCCGCGG ACCCCTCCCAACCCTGCCTTTGCGGCCCGGCACATTAGGAAATACTCCCC		
	ORF Start: ATG at 2	ORF Stop: TGA at 1757	
	SEQ ID NO: 32	585 aa	MW at 65861.1kD
NOV12a, CG100710-01 Protein Sequence	MSWLFQVNGKPGKGGAGPPPLPPAQPAGGEGDRGLGDRPAPKDKWSNFDPTGLERA AKAARELEHSRQTQERLSGSASPEQVPAGECCALQEYEAQVQLKSEQIRAEERRK TLSEETRQHQARAQYQDKLARQRYEDQLKQQQLLNEENLRKQEEVSQKQAMRRATVE REMELRHKNEMLRVETEARARAKAERENADIIREQIRLKASEHRQTVLESIRTAGTLF GEGFRAFVTDRDKVTATVAGLTLLAVGVYSAKNATAVTGRFIEARLGKPSLVRETSRI TVLEALRHPIQVSRRLSRPDVLEGVVLSPLSLEARVRDIAIATRNTKKNRGLYRHIL LYGPPGTGKTLFAKKLALHSGMDYAIMTGGDVAPMGREGVTAMHKLFDWANTSRRGLL LFMDEADAFLRKRATEEISKDLRATLNAFLYHMGQHSNKFMLVLASNLPEQFDCAINS RIDVMVHFDLPQEEERLVRHLFDNCVLKPATEGKRRLLKLAQFDYGRKCEVARLTE MSGREIAQLAVSWQATAYASEDGLTEAMMDTRVQDAVQQHQMMRWLKRGRPGPED EQPSS		

Further analysis of the NOV12a protein yielded the following properties shown in Table 12B.

Table 12B. Protein Sequence Properties NOV12a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1524 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV12a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 12C.

Table 12C. Geneseq Results for NOV12a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV12a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB93086	Human protein sequence SEQ ID NO:11926 - Homo sapiens, 586 aa. [EP1074617-A2, 07-FEB-2001]	1..585 1..586	529/586 (90%) 547/586 (93%)	0.0
AAM39035	Human polypeptide SEQ ID NO 2180 - Homo sapiens, 602 aa. [WO200153312-A1, 26-JUL-2001]	70..585 25..540	502/516 (97%) 505/516 (97%)	0.0
AAB94853	Human protein sequence SEQ ID NO:16038 - Homo sapiens, 480 aa. [EP1074617-A2, 07-FEB-2001]	168..585 1..418	404/418 (96%) 407/418 (96%)	0.0
ABB71478	Drosophila melanogaster polypeptide SEQ ID NO 41226 - Drosophila melanogaster, 599 aa. [WO200171042-A2, 27-SEP-2001]	1..570 1..576	342/579 (59%) 423/579 (72%)	0.0
AAB63264	Human breast cancer associated antigen protein sequence SEQ ID NO:626 - Homo sapiens, 208 aa. [WO200073801-A2, 07-DEC-2000]	279..474 3..198	181/196 (92%) 187/196 (95%)	8e-99

In a BLAST search of public sequence databases, the NOV12a protein was found to have homology to the proteins shown in the BLASTP data in Table 12D.

Table 12D. Public BLASTP Results for NOV12a				
Protein Accession Number	Protein/Organism/Length	NOV12a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BUK4	SIMILAR TO HYPOTHETICAL PROTEIN FLJ10709 (HYPOTHETICAL 72.6 KDA PROTEIN) - Homo sapiens (Human), 648 aa.	1..585 1..586	551/586 (94%) 561/586 (95%)	0.0
Q9NVI7	CDNA FLJ10709 FIS, CLONE NT2RP3000869 - Homo sapiens (Human), 586 aa.	1..585 1..586	529/586 (90%) 547/586 (93%)	0.0
Q96A50	HYPOTHETICAL 66.3 KDA PROTEIN (UNKNOWN) (PROTEIN FOR MGC:14291) (PROTEIN FOR MGC:20264) - Homo sapiens (Human), 586 aa.	1..585 1..586	528/586 (90%) 546/586 (93%)	0.0
Q96T67	TOB3 - Homo sapiens (Human), 578 aa.	1..572 1..572	527/573 (91%) 541/573 (93%)	0.0
Q92511	TOB3 - Mus musculus (Mouse), 591 aa.	1..583 1..583	500/584 (85%) 540/584 (91%)	0.0

PFam analysis predicts that the NOV12a protein contains the domains shown in the Table 12E.

Table 12E. Domain Analysis of NOV12a			
Pfam Domain	NOV12a Match Region	Identities/ Similarities for the Matched Region	Expect Value
AAA	346..542	60/221 (27%) 140/221 (63%)	2.6e-32

Example 13.

- 5 The NOV13 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 13A.

Table 13A. NOV13 Sequence Analysis			
	SEQ ID NO: 33	1705 bp	
NOV13a, CG100730-01 DNA Sequence	GATAAATGGAGGAAGCTAAGACCTCACTGAGCCTGTGGCTCCGGCCGCTGAGTGATG TCCCTTCTGTCTGCCAGGGTGGGATCGGGCACTCATCCAGTTGCAGGTGCGAGCACT TTGAGTAGCACAGGGTCTCAAGCTGTGGCCATGGACTTGGGCAACAGGAAATTAGAA TATCCTCTGGAAAAGCTGGCCAGATTGCAAAATGGCTCTGCTGTAATACAGTCAGGTG CACTGCGGTAATGGCCACAGCCGTCAGTAAACAAAACCTTCCCCTTCCCAGTTTATG CCTTTGGTGGTTGACTACGGACAAAAGGCTGCTGCAGTAGGTAGAATCCCACAAACC CTCTGAGAAGAGAGATTAGTACTTCTGATAAAGAAATCTTACACGTGGAATAGTAG TTGTTCAAGTTAGACCTATCTTCCAGCTGGCTACTTCTGTGATATACAGGTTCTGTG AATCTGTTAGCAGTAGATGGTATTAAATGAACTTGATGTCTAGCAATTAATGGCAACT CTGTAGCCCTCTCATTATCAGATATTCTTTGGAATGGACCTGTTGGGACAGTAAGAAT AGGAATGACTGATGGAGAATGTGTTGTTAACCACAAGAAAAGAAATGTCTTCTAGT ACTTTAAATGTAGTGGTTGCTGGAGCACCTAAAAGTCAGACTGTCATGTTGGAAGCCT CTGCAGAGAACATTTTACAGCAGGACTTTTGCCACGCTATCAAAGTGGGAGTGAATA TACCAGCAAATAATTACGGGCATCCAGCAGCTGGTAAAAGAAATGGTGTACCAAG AGGCACCTCAGAAGTTATTTACCCCTTCACAAGAGATTGTGAAACATGCTCATAAAG TCATATGGAGAGACTCTATGCAGTTTTACAGATTATGAACATGATAAAATTTCCAG AGATGAAGCTGTTAAACAAAATAAGATTAGATACAGAGGAACAACTAAAAGAAATATT CCAGAGGTCGATCTATATGAGATAATAGAATCCTTCAATACTGTTGCAAGAAAGTTT TTAGAAGTATTATTTGAATGAATACAAAAGATGCGATGGTCGGGATTTGACTTTACT TCGGAATATAAGTTGTGAGGTAGATATGTTTAAACCCTTCATGGATCAGAAATATT CAAAGAGGACAAACACAGCTGCTTTGTGCTGTTACATTGATTCAATAGAAATCCAGTA TTAAGTTGGATCGAGTTATAACAACTATAAATGGGATAAAAGATAAAATTTTCATGCT GCACTATGAGTTTCTCCTTATGCAACTAATGAAATGTCAAAGTCACCGGTATAAAT AGAAGAGAACTTGGGCCTGATGCTCTTGCTGAGAAAGCTTTGTATCCTGTTATCCCC AAGATTTTCTTTTGCCATAAGAGTTACATCTGAAGTCCTAGAGTCAAATGGGTCATC TTCTATGGCATCTGCATGTGGTGAAGTTTAGCATTAAATGGATTAGGGCTTCCAATT AAATCTGCTGTACAGGTGTAGCAATGGGATTGGCCACCAAAACCGATCTTGAGAAGG GTGAAATAGAAGATTATCATTTGCTAACAGATATTCTGGGAATTGAAGCTTACAATGG TTGACATGGACTTCAAATAGCTGGCACTAATAAAGGAATAACTGCATTACAGGCAGA TATTAAATTACCTGGAATAACAA		
	ORF Start: ATG at 6	ORF Stop: TGA at 1626	
	SEQ ID NO: 34	540 aa	MW at 59197.4kD
NOV13a, CG100730-01 Protein Sequence	MEEAKTSLSLWLRPLSDGPFCLPGWDRALIQLQVRALLSSTGSQAVAMDIGNRKLEIS SGKLARFANGSAVIQSGDTAVMATAVSKTKPSPSQFMPLVVDYGGQKAAAVGRIPTNPL RREISTSDKEILTRGIVDCSVRPIFPAGYFCDIQVLCNLLAVDGINELDLVLAINGNSV ALSLSDILWNGPVGTVRIGMTDGECEVNPTRKEMSSSTLNVVAGAPKQVTVMLEASA ENILQQDFCHAIKVGKVTQQIIQGIQQLVKEIGVTKRTPQKLFPSQBIKHAHKL MERLYAVFTDYEHDKISRDEAVNKIRLDTEEQLKEIFPEVDLYEIIESFNTVAKKVF SIILNEYKRCDDRLTLRLNISCEVDMFKTLHGSELFQRGQTQLCAVTFDSLESSIK LDRVITTINGIKDKNFMLHYEFPYPATNEIVKVTGINRRELGPDALAEKALYVPIPKD FPFAIRVTSEVLESNGSSMASACGSLALMDSGLPIKSAVTGVAMGLATKTDLEKGE IEDYHLLTDILGIEAYNG		

Further analysis of the NOV13a protein yielded the following properties shown in Table 13B.

Table 13B. Protein Sequence Properties NOV13a	
PSort analysis:	0.6500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen); 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV13a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 13C.

Table 13C. Geneseq Results for NOV13a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV13a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB08732	Amino acid sequence of a human OLD-35 polypeptide - Homo sapiens, 705 aa. [WO200046391-A2, 10-AUG-2000]	17..540 1..524	455/524 (86%) 477/524 (90%)	0.0
AAB92684	Human protein sequence SEQ ID NO:11065 - Homo sapiens, 504 aa. [EP1074617-A2, 07-FEB-2001]	10..466 11..467	396/457 (86%) 416/457 (90%)	0.0
ABG17275	Novel human diagnostic protein #17266 - Homo sapiens, 899 aa. [WO200175067-A2, 11-OCT-2001]	192..540 522..852	328/349 (93%) 330/349 (93%)	0.0
ABG08546	Novel human diagnostic protein #8537 - Homo sapiens, 899 aa. [WO200175067-A2, 11-OCT-2001]	192..540 522..852	328/349 (93%) 330/349 (93%)	0.0
ABG17275	Novel human diagnostic protein #17266 - Homo sapiens, 899 aa. [WO200175067-A2, 11-OCT-2001]	192..540 522..852	328/349 (93%) 330/349 (93%)	0.0

In a BLAST search of public sequence databases, the NOV13a protein was found to have homology to the proteins shown in the BLASTP data in Table 13D.

Table 13D. Public BLASTP Results for NOV13a				
Protein Accession Number	Protein/Organism/Length	NOV13a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9DC52	I200003F12RIK PROTEIN - Mus musculus (Mouse), 540 aa.	12..533 13..534	434/522 (83%) 467/522 (89%)	0.0
Q96T05	CDNA FLJ14531 FIS, CLONE NT2RM2000371, WEAKLY SIMILAR TO POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (EC 2.7.7.8) - Homo sapiens (Human), 504 aa.	10..466 11..467	396/457 (86%) 416/457 (90%)	0.0
Q9V9X7	CG11337 PROTEIN - Drosophila melanogaster (Fruit fly), 748 aa.	35..540 29..540	262/512 (51%) 354/512 (68%)	e-143
Q95RX7	LD03255P - Drosophila melanogaster (Fruit fly), 720 aa.	55..540 1..489	257/489 (52%) 342/489 (69%)	e-142
Q9S7G6	POLYNUCLEOTIDE PHOSPHORYLASE - Arabidopsis thaliana (Mouse-ear cress), 991 aa.	49..540 58..545	201/494 (40%) 307/494 (61%)	e-104

PFam analysis predicts that the NOV13a protein contains the domains shown in the Table 13E.

Table 13E. Domain Analysis of NOV13a			
Pfam Domain	NOV13a Match Region	Identities/ Similarities for the Matched Region	Expect Value
RNase_PH	48..250	58/255 (23%) 137/255 (54%)	4.5e-20
RNase_PH	357..537	58/260 (22%) 136/260 (52%)	3.2e-15

Example 14.

- 5 The NOV14 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 14A.

Table 14A. NOV14 Sequence Analysis

Table 14A. NOV14 Sequence Analysis			
	SEQ ID NO: 35	2626 bp	
NOV14a, CG100819-01 DNA Sequence	CCTGCGTTGACCGCGTGC CGGGTGT CATGGCGGCTGCAGGTACTGCTGCTCGTGCCT CCGGCTCCGGCCCCCTGAGCGATGGTCCCTTTCCTTCTGCCACGGCGGGATCGGGCACT ACCCAGTTGCAAGTGCAGGCACTATGGAGTAGCGCAGGGTCTCGAGCTGTGGCCGTGG ACTTAGCCAAACAGGAAATTAGAAATATCTTCTGGAAGCTGGCCAGATTGTCAGATGG CTCTGCTGTAGTACAGTCAAGTGACACTGCAGTAATGGTCACAGCGGTCACTAAACA AAACCTTCCCTTCCCAGTTTATGCCTTTGGTGGTAGACTACAGACAAAAGCTGCTG CAGCAGGTAGAATTCCCACAACTATCTGAGAAGAGAGATTGGTACTTCTGATAAAG AATTCTAACCAAGTCGAATAATAGATCGTTCAATTAGACCGCTCTTCCAGCTGGCTAC TTCTATGATACACAGGTACTGTGAATCTGTTAGCAGTAGATGGTGAATAGAGCCTG ATGTCCTAGCAATTAATGGCGCTTCCGTAGCCCTCTCATTATCAGATATTCCCTGGAA TGGACCTGTTGGTGCAGTACGAATAGGAATAATTGATGGAGAATATGTTGTTAAACCC ACAAGAAAAGAAATGTCTTCTAGTACTTTAAATTTAGTGGTGGTGGAGCACCTAA GTCAGATTGTCTGTTGGAAGCCTCTGCAGAGAACATTTTACAGCAGGACTTTTGCC TGCTATCAAAGTGGGAGTGAAATATACCCAACAAATAATTCAAGGCATTCAAGAGTT GTAAAGAACTGGTGTACCAAGAGGACACCTCAGAAGTTATTACCCCTTCGCCAG AGATTGTGAAATATACTCATAAACTTGCTATGGAGAGACTCTATGCAGTTTACAGA TTACGAGCATGACAAAGTTTCCAGAGATGAAGCTGTTAACAAATAAGATTAGATACG GAGGAACAACTAAAAGAAAAATTTCCAGAAGCCGATCCATATGAATAATAGAATCCT TCAATGTTGTTGCAAGGAAGTTTGTAGAAGTATTGTTTGAATGAATACAAAGGTTG CGATGGTCCGGATTGACTTCACTTAGGAATGTAAGTTGTGAGGTAGATATGTTTAA ACCCCTTCATGGATCAGCATTATTTCAAAGAGGACAAACACAGGTGCTTTGTACCGTTA CATTGATTCAATAGAACTCTGGTATTAAAGTCAGATCAAGTTATAACAGCTATAAATGG GATAAAAGATAAAAAATTTATGCTGCACTACGAGTTTCTCTTATGCAACTAATGA ATTGGCAAAGTCACTGGTTTAAATAGAAGAGAACTTGGGCATGGTGCTCTTGCTGAGA AAGCTTTGTATCCTGTTATTTCCAGAGATTTTCTTTCCACATAAGAGTTACATCTGA AGTCTAGAGTCAAATGGGTCACTCTTATGGCATCTGCATGTGGCGGAAGTTTAGCA TTAATGGATTCAAGGGTTCCAAATTTATCTGCTGTTGCGAGCGTAGCAATAGGATTGG TCACCAAAACCGATCCTGAGAGGGTGAAATAGAAGATTATCGTTTGCTGACAGATAT TTTGGGAATTGAAGATTACAATGGTGACATGGACTTCAAATAGCTGGCACTAATAAA GGAATAACTGCATTACAGGCTGATATTAATTAACCTGGAATACCAATAAAAAATTGTGA TGGAGGCTATTCAACAAGCTTCAGTGGCAAAAAGGAGATATTACAGATCATGAACAA AACTATTTCAAACCTCGAGCATCTAGAAAAGAAAATGGACCTGTTGTAGAACTGTT CAGGTTCCATTATCAAAACGAGCAAAATTTGTTGGACCTGGTGGCTATAACTTAAAA AACTTCAGGCTGAAACAGGTGTAATATTAGTCAGGTGGATGAAGAAACGTTTTCTGT ATTGCAACCAACACCCAGTGTATGTCATGAGGCAAGAGACTTCATTACTGAAATCTGC AAGGATGATCAGGAGCAGCAATTAGAATTTGGAGCAGTATATACCGCCACAATACTG AAATCAGAGATACTGGTGAATGGTAAATATATCCAATATGACTCGGGTACTGCT TCATAACACACAACCTTGATAACGAAAGATTAAACATCCTACTGCCCTAGGATTAGAAG TTGGCCAAAGAAATTCAGGTGAAATACTTTGGACGTGACCCAGCCGATGGAAGAATGAG GCTTCTCGAAAAGTGCTTCAGTCGCCAGCTACAACCGTGCTGAGAATTTGAATGAC AGAAGTAGTATTGTAATGGGAGAACCTATTTACAGTCATCATCTAATTTCTAGTAGT TTTTTTTTTTTAAAGAGAATTTCTAGAATTTCTATTTGTCTAGGGTGATGTGCTGTAGA GCAACATTTTAGTAGATCTTCCATTGTGTAGATTCTATATAATAATAATACATTTTA ATTATTGTACTAAAAATGCTCATTTACATGTGCCATTTTTTAAATTCGAGTAACCCAT ATTTGTTAATTTGATTATTACATTATAAATCAAGAAATATTTATTATTAAGTAAGTC ATTTATACATCTTAGA		
	ORF Start: ATG at 27	ORF Stop: TAG at 2193	
	SEQ ID NO: 36	722 aa	MW at 79307.2kD
NOV14a, CG100819-01 Protein Sequence	MAACRYCCSCLRLRPLSDGPFLLPRDRALTQLQVRALWSSAGSRVAVDLGNRLKLEI SSGKLARFADGSVAVQSGDTAVMVTAVSKTKPSQFPLVVDYRQKAAAAGRIPTNY LRREIGTSDKEILTSRIIDRSIRPLFPAGYFYDTQVLCNLLAVDGVNEPDLAINGAS VALSLSDIPWNGPVGAVRIGIIDGEYVNPTRKEMSSSTLNLVVAGAPKSQIVMLEAS AENILQQDFCHAIKVGKYTQQIIQGIQQLVKETGVTKRTPQKLFPTSPFIVKTHKL AMERLYAVFTDYEHDKVSRDEAVNKIRLDTEEQLEKFPFADPYEIIIESFNVAKEVF RSIVLNEYKRCRDLTSLRNVSCVDMFKTLHGSALFQRGQTQVLCVTTFDSLESIGI KSDQVITAINGIKDKNFMLHYEFPYPATNEIGKVTGLNRRLEHGHALAEKALYPVPIPR DFPFTIRVTSEVLESNGSSMASACGGSLALMDSGVPISSAVAGVAIGLVTKDPEKG EIEDYRLRLTDILGIEDYNGDMDFKIAGTNKGITALQADIKLPGIPKIVMEAIQQASV AKKEILQIMNKTIKPRASRKENGFPVETVQVPLSKRAKFPVGGYNLKKLQAGTGVV ISQVDEETFSVFAPTPSVMHEARDFITEICKDDQEQQLFEGAVYTTATITEIRDTGVMV KLYPNMTAVLLHNTQLDNERLNILLP		

Further analysis of the NOV14a protein yielded the following properties shown in Table 14B.

Table 14B. Protein Sequence Properties NOV14a	
PSort analysis:	0.5016 probability located in mitochondrial matrix space; 0.4500 probability located in cytoplasm; 0.2212 probability located in mitochondrial inner membrane; 0.2212 probability located in mitochondrial intermembrane space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV14a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 14C.

Table 14C. Geneseq Results for NOV14a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV14a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB08732	Amino acid sequence of a human OLD-35 polypeptide - Homo sapiens, 705 aa. [WO200046391-A2, 10-AUG-2000]	18..722 1..705	704/705 (99%) 705/705 (99%)	0.0
AAB92684	Human protein sequence SEQ ID NO:11065 - Homo sapiens, 504 aa. [EP1074617-A2, 07-FEB-2001]	1..467 1..467	466/467 (99%) 467/467 (99%)	0.0
ABB58546	Drosophila melanogaster polypeptide SEQ ID NO 2430 - Drosophila melanogaster, 748 aa. [WO200171042-A2, 27-SEP-2001]	36..717 29..693	363/688 (52%) 485/688 (69%)	0.0
ABG17275	Novel human diagnostic protein #17266 - Homo sapiens, 899 aa. [WO200175067-A2, 11-OCT-2001]	190..583 519..894	335/394 (85%) 349/394 (88%)	0.0
ABG08546	Novel human diagnostic protein #8537 - Homo sapiens, 899 aa. [WO200175067-A2, 11-OCT-2001]	190..583 519..894	335/394 (85%) 349/394 (88%)	0.0

In a BLAST search of public sequence databases, the NOV14a protein was found to have homology to the proteins shown in the BLASTP data in Table 14D.

Table 14D. Public BLASTP Results for NOV14a				
Protein Accession Number	Protein/Organism/Length	NOV14a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9DC52	I200003F12RIK PROTEIN - Mus musculus (Mouse), 540 aa.	1..534 1..534	478/534 (89%) 506/534 (94%)	0.0
Q96T05	CDNA FLJ14531 FIS, CLONE NT2RM2000371, WEAKLY SIMILAR TO POLYRIBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (EC 2.7.7.8) - Homo sapiens (Human), 504 aa.	1..467 1..467	466/467 (99%) 467/467 (99%)	0.0
Q95RX7	LD03255P - Drosophila melanogaster (Fruit fly), 720 aa.	56..717 1..665	373/665 (56%) 494/665 (74%)	0.0
Q9V9X7	CG11337 PROTEIN - Drosophila melanogaster (Fruit fly), 748 aa.	36..717 29..693	363/688 (52%) 485/688 (69%)	0.0
Q9S7G6	POLYNUCLEOTIDE PHOSPHORYLASE - Arabidopsis thaliana (Mouse-ear cress), 991 aa.	50..715 58..713	278/668 (41%) 417/668 (61%)	e-148

PFam analysis predicts that the NOV14a protein contains the domains shown in the Table 14E.

Table 14E. Domain Analysis of NOV14a			
Pfam Domain	NOV14a Match Region	Identities/ Similarities for the Matched Region	Expect Value
RNase_PH	48..251	68/255 (27%) 147/255 (58%)	1.1e-37
RNase_PH	358..581	75/261 (29%) 175/261 (67%)	4e-55
KH-domain	609..651	14/49 (29%) 31/49 (63%)	0.35

Example 15.

- 5 The NOV15 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 15A.

Table 15A. NOV15 Sequence Analysis

	SEQ ID NO: 37	2030 bp	
NOV15a, CG100872-01 DNA Sequence	CCATGTCCTTCCAGAGTATCATCCACCTGTCCCTGGACAGCCCTGTCCATGCCGTTTCTGTGTTGGGCACAGAAATCTGCTTGGATCTCACCAGGTGTGCCCCCAGAAAGTGCCAGTGCTTCACCATCCATGGCTCTGGGAGGGTCTTGATCGATGTGGCCAAACACGGTGATTCTGAGAAGGAGGACGCCACCATCTGGTGGCCCTGTCTGATCCACGTACGCCACAGTGAAGATGACATCGCCAGCCCTTCCGTGGATGCGGATAAGGTAAGCCTCACATACTATGGGCCCCAACGAGGATGCCCCCGTGGGCACAGCTGTGCTGTACCTCACTGGCATTGGTGAGTGTGCTCCAACCTGGGAGCCTCCCAACCGCTGGGACCCAGATCCATTTTTTCTCAAAAAATGGATCTGGGGTCCCAGCGTTGGGGTGCCATCTGCTTGTGAATTGCAACCTCTGCTGATGTGGGCCAGCAACTTGAGGACAAGAAAAACCAAGAAAGTGATCTTTTCAGAGGAAATAACGAATCTGTCCAGATGACTCTGAATGTCCAAGGCCACAGCTGTATCTTTAAAGAAATATCGGCTAGTCTCTCCATACCTCCAAGGAAGAGTCGAAGAAGGCGAGAGTCTACTGGCCCCAAAGTGAAGACAATCCAGTACCTTTGAGTTGGTGTGGGGCCCCGACAGCACGCCTTATACCTTGGCCCTCTCGGGGAACCACTTGAAGGAGACTTTCTACGTTGAAGCTATAGCATTTCCCATCTGCCGAATTCTCAGGCCTCATCTCTACTCTGTGTCCCTGTGGGAGGAGTCTTTGCACTCAATTCAGAGACTGTGCTGTACAAAGACACGGTGGTGTGTCGAGTGTGCTCCCTGTGTCTTCACTCCCTGTACCCAGGTGCCTCTGGAGGTTTACCTGTGCTATCTGAGCTGGGTGTTGAGGGTTGAGCAGGAGTTGGCAGGAAGGAAGGGGGTTCTTACCGTACTTTTCCGGCAACAATCTGTCTTCTCTCCATTCCAGGATGAGATGGCCTTCTGCTACACCCAGGCTCCCCACAAGACAACGTCCTTGATCCTCGACACACCTCAAGGCCGCCGATCTCGATGAGTTCCCCATGAAGTACTCACTGGTGCCTGGTATTGGCTACATGATCCAGGACACTGAGGACCATAAAGTGGCCAGCATGGATTCCATTGGGAACCTGATGGTGTCCCCACCTGTCAAGGTCCAAGGGAAGAGTACCCGCTGGGACAGATCCTCATTGGCAGCAGCTTTTACCCCTCTAGCGCAGAGGGCCGGGCCATGAGTAAGACCCTCCGAGACTTCTCTATGCCCAGCAGGTCCAAGCGCCGGTGGAGCTCTACTCAGATTGGCTAATGACTGGCCACGTGGATGAGTTTATGTGCTTATCCCCACAGATGACAAGCAGGGCTTCTGTGCTCTCTGGCCAGCCCCAGTGCCTGCTATAAACTGTTCCGAGAGAAACAGAAGGAAGGCTATGGCGACGCTCTTCTGTTTGATGAGCTTAGAGCAGATCAGCTCCTGTCTATGGTAAGGGAATCCCTTTGCCCCCCCCCCCCACCCACCCACCCACCAAGAAGTGCACTTCACTGAACCCGTGACATCCTGAAGACGGAGCTGGGCTGGTGGAAACAGGACATCATCGAGATTCCCCAGCTGTTCTGCTTGGAGAAGCTGACTAACATCCCCTCTGACCAGCAGCCCAAGTTGCGGATGATTGTGATGGGCAAGAACCTGGGGATCCCCAAGCCTTTTGGGCCCAAAATCAAGGGGACCTGCTGCTGGAAGAAAAGATTGCTGCTTGTCTGGAGCCCCTGGGCTTCAAGTGCACCTTCATCAATGACTTTGACTGTTACTGACAGAGGTGGAGACATCTGTGCTGTGCCAATCCGCGGGGTGCCCTTTGCCTTCAAAATGGTGGAAGATGGTACCTTAGACCCAGGCCCTGGAGCTGCCAGCTCTGCCCCAGCGTGGATGGCCCACT		
	ORF Start: ATG at 3	ORF Stop: TAG at 1980	
	SEQ ID NO: 38	659 aa	MW at 73697.2kD
NOV15a, CG100872-01 Protein Sequence	MSFQSI IHLSDSPVHAVCVLGTETCLDLTRCAPQKQCFTIHGSGRVLIDVANTVIS EKEDATIWWPLSDPTYATVKMTSPSPSVDADKVSLEYGPNEDAPVGTAVLYLTGIGECASNWEPPNRWDPRSIFSQKKWIWGPSGWGAILLVNCPADVGGQLEDKTKKVFSEETITNLSQMTLNVQGPSKILKKYRLVLHTSKEESKKARVYWPQSEDNSSTFELVLGPDQHAYTLALLGNHLKETFYVEAIAFPSAEFSGLISYSVSLVEESLQSIPTVLYKDTVVRVAPCVFIPCTQVPLEVYLCYLSWVLRVEQELAGRKGVLTVPFPANNLSSSPQDEMAFCYTQAPHKTTSLIDTPQAADLDEFPMKYSLVPGIGYMIQDTEHDKVASMDSIGNLMVSPPVKVQKEYPLGRVLIGSSFYPSAEGRAMSKTLRDFLYAQVQAPVELYSDWLMTGHVDEFMCFIPTDDKQGFLLLASPSACYKLFREKQKEGYDALLFDELRADQLLSNKGKTPFRPPPHPPHPQKCIHLNRDILKTELGLVEQDIIETIPQLFCLEKLTNIPSDQQPKLRMIVMGKNLGIKPFPGQIKGTCCLEEKICLLEPLGFKCTFINDFDCYLTEVGDICACANIRRVPPFAFKWKMVP		

Further analysis of the NOV15a protein yielded the following properties shown in Table 15B.

Table 15B. Protein Sequence Properties NOV15a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3600 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen); 0.0598 probability located in microbody (peroxisome)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV15a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 15C.

Table 15C. Geneseq Results for NOV15a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV15a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE06034	Mouse oocyte protein 5 (MOP5) - Mus musculus, 664 aa. [WO200153339-A2, 26-JUL-2001]	20..659 1..664	386/672 (57%) 492/672 (72%)	0.0
AAB41584	Human ORFX ORF1348 polypeptide sequence SEQ ID NO:2696 - Homo sapiens, 663 aa. [WO200058473-A2, 05-OCT-2000]	1..659 1..663	289/666 (43%) 411/666 (61%)	e-155
ABB57299	Mouse ischaemic condition related protein sequence SEQ ID NO:839 - Mus musculus, 673 aa. [WO200188188-A2, 22-NOV-2001]	7..659 15..673	268/667 (40%) 379/667 (56%)	e-132
ABG26518	Novel human diagnostic protein #26509 - Homo sapiens, 761 aa. [WO200175067-A2, 11-OCT-2001]	31..632 36..642	263/620 (42%) 361/620 (57%)	e-122
ABG26518	Novel human diagnostic protein #26509 - Homo sapiens, 761 aa. [WO200175067-A2, 11-OCT-2001]	31..632 36..642	263/620 (42%) 361/620 (57%)	e-122

- 5 In a BLAST search of public sequence databases, the NOV15a protein was found to have homology to the proteins shown in the BLASTP data in Table 15D.

Table 15D. Public BLASTP Results for NOV15a				
Protein Accession Number	Protein/Organism/Length	NOV15a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9UM07	Protein-arginine deiminase type V (EC 3.5.3.15) (Peptidylarginine deiminase V) (HL-60 PAD) - Homo sapiens (Human), 663 aa.	1..659 1..663	289/666 (43%) 411/666 (61%)	e-154
O88807	Protein-arginine deiminase type IV (EC 3.5.3.15) (Peptidylarginine deiminase IV) (PAD-R4) (Peptidylarginine deiminase type alpha) - Rattus norvegicus (Rat), 666 aa.	1..659 1..666	297/675 (44%) 404/675 (59%)	e-148
P70708	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Rattus norvegicus (Rat), 664 aa.	1..659 1..664	290/677 (42%) 408/677 (59%)	e-145
Q9Z184	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Mus musculus (Mouse), 664 aa.	1..659 1..664	288/677 (42%) 409/677 (59%)	e-144
O57400	PEPTIDYLARGININE DEIMINASE - Gallus gallus (Chicken), 672 aa.	1..659 1..672	286/682 (41%) 399/682 (57%)	e-143

PFam analysis predicts that the NOV15a protein contains the domains shown in the Table 15E.

Table 15E. Domain Analysis of NOV15a			
Pfam Domain	NOV15a Match Region	Identities/ Similarities for the Matched Region	Expect Value
PAD	1..659	311/702 (44%) 446/702 (64%)	4.7e-222

Example 16.

- 5 The NOV16 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 16A.

Table 16A. NOV16 Sequence Analysis			
	SEQ ID NO: 39	2018 bp	
NOV16a, CG100980-01 DNA Sequence	GATGGCCCCAAGAGAGTTGTGTCAGCTGTCCCTGAAGATGCCTACCCATGCCGTGTGT GTGGTGGGAGTCGAGGCACATGTGGACATTACAGTGTATGTGCCAAGGGTGCCAACA GCTTCAGGGTCTCTGGAAGCTCCGGGGTGGAGGTCTTTCATGGTCTACAACCGCACACG TGTGAAAGAGCCCATAGGCAAGGCCGTTGGCCGCTAGACACTGATGCAGACATGGTC GTATCTGTGGGCACAGCCAGTAAGGAATTAAAGGACTTCAAGGTAAGAGTCTCCTACT TTGGGGAGCAGGAAGACCAAGCTCTGGGCCGCAGCGTGCTTTACCTCACTGGCGTCTCT AGAACCCCTTCTGGAGCCTCTGGGTCCCACTGATCCCAAGGCATCTTATTTGGCCA CAGAAAACCTGGCGCTGGGGCCCTGAGGGCTATGGGGCTATCTTGTGCTGTAAGTGTG ACCGGGACAATCACAGGTCCGCAGAGCCTGACCTCACCACAGCTGGCTGATGTGCGT GGCTGACCTGCAGGACATGTCCCCAATGCTGCTGAGCTGCAATGGCCCCGACAAGCTC TTGACAGCCACAAGCTTGTCTTGAACGTGCCCTTTTCTGATTCCAAAAGAGTGAAGG TCTTCTGTGCCAGGGGTCTGAGGATGTGTGTGAGGCCTATAGGCATGTGCTGGGCCA AAATAAGGTGTCTATGAGGTACCCCGCTTGCATGGGGATGAGGAGCGCTTCTTCGTG GAAGGCCTGTCTTCCCTGATGCCGGCTTACAGGACTCATCTCTTCCATGTCACTC TGCTGGACGACTCCAACGAGGATTTCTCGGCATCCCCTATCTTCAAGACACTGTGGT GTTCCGAGTGGCACCCTGGATCATGACGCCAGCACTCTGCCACCCCTAGAGGTGTAT GTGTGCCGGGTGAGGAACAACAGCTGTTTGTGGATGCGGTGGCAGAGCTGGCCAGGA AGGCCGGCTGCAAGCTGACCATCTGCCACAGGCCGAGAACCGCAACGACCGCTGGAT CCAGGATGAGATGGAGCTGGGCTACGTTTCAAGCGCCGACACAAGACCTTCCGGTGGT TTTGACTCCCCAAGGAATGGGGAATGCAGGATTTCCCTTACAAAAGAAATCTGGGTG CAGATTTTGGTTACGTGACTCGGGAACACGCGACAGGTCTGTGAGTGGCCTGGACTC CTTTGGGAACCTGGAGGTGAGCCCTCCAGTGGTGGCCAAATGGGAAGAGTACCCCTG GGGAGGATCCTCATTGGGGGCAACCTGCCTGGGTCAAGTGGCCGAGGGTCACCCAGG TGGTGGCGGACTTCCTCCATGCCAGAAAGTGCAGCCCCCGTGGAGCTCTTTGTGGA CTGGTTGGCCGTGGGCCATGTGGATGAGTTTCTGAGCTTTGTCTCCTGCCCGATGGG AAGGGCTTCCGGATGTCTCTGGCCAGCCCTGGGGCTGCTTCAAGCTCTTCCAGGAAG AGCAGAAGTGTGCCACGGGAGGGCCCTCCTGTTCCAGGGGGTGTGTTGGTGTATGAGC GGTCAAGACCATCTCCATCAACCAGGTGCTCTCCAATAAAGACCTCATCAACTACAAT AAGTTTGTGCAGAGCTGCATTGACTGGAACCGTGAGGTGCTGAAAGCGGAGCTGGGCC TGGCAGAGTGTGACATCATTGACATCCCAAGCTCTTCAAGACCGAGAGGAAAAAGC AACGGCCTTCTTCCCTGACTTGGTGAACATGCTGGTGTCTGGGGAAGCACTGGGCAT CCCAAGCCCTTTGGGCCCATCATCAATGGCTGTGCTGCTGGAGAGGAGAAGTGGCGT CCCTGCTGGAGCCGCTGGGCCCTCCACTGCACCTTCATTGATGACTTCACTCCATACCA CATGCTGCATGGGGAGGTGCCTGTGGCACCAATGTGTGCAGAAAGCCCTTCTCTTTT AAGTGGTGAACATGGTGCCTGAGACAGCTCCCACCACCATCT		
	ORF Start: ATG at 2	ORF Stop: TGA at 1994	
	SEQ ID NO: 40	664 aa	MW at 74356.9kD
NOV16a, CG100980-01 Protein Sequence	MAPKRVVQLSLKMPHVAVCVVGVGEAHVDIHS DVPKGANSFRVSGSSGVEVFMVYNRT VKEPIGKARWPLDADMDVSVGTASKELKDFKVRVSYFGEQEDQALGRSVLYLTGVL EPLPGASWVPLIPRHLILPQKTWRWGPEGYGAILLVNCDRDNHRSAPEDLTHSWLMS ADLQDMSPLLSCNGPDKLFD SHKLVLNVFPDSKRVRVFCARGPEDVCEAYRHVLGQ NKVSYEVPRLHGDEERFFVEGLSFPDAGFTGLISFHVTLLDSDNEDFSASPIPTDTV FRVAPWIMTPSTLPPLEVYVCRVRNNTCFVDAVAELARKAGCKLTICPOAENRNRDW QDEMELGYVQAPHKTLPPVFDSPRNGELQDFPYKRIILGPDFGYVTREPRDRSVSGLD FGNLEVPVAVANGKEYPLGRILIGGNLPGSSGRRVTQVVRDFLHAQKVQPPVFLFV WLAVGHVDEFLSFPAPDGKGFRLMLASPGACFKLFQEKQKCGHGRALLFQGVGDE VKTISINQVLSNKDLINYNKFVQSCIDWNREVLKRELGLAECDIIDIPLQPKTERKK TAFPPDLVNMLVLGKHLGIPKPPGPIINGCCLEEKVRSILLEPLGLHCTFIDDFTPY MLHGEVHCNTNVCRKPFSPKWNMVP		

Further analysis of the NOV16a protein yielded the following properties shown in Table 16B.

Table 16B. Protein Sequence Properties NOV16a	
Psort analysis:	0.5500 probability located in endoplasmic reticulum (membrane); 0.2424 probability located in lysosome (lumen); 0.1410 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV16a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 16C.

Table 16C. Geneseq Results for NOV16a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV16a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG26518	Novel human diagnostic protein #26509 - Homo sapiens, 761 aa. [WO200175067-A2, 11-OCT-2001]	32..645 37..650	591/614 (96%) 595/614 (96%)	0.0
ABG26518	Novel human diagnostic protein #26509 - Homo sapiens, 761 aa. [WO200175067-A2, 11-OCT-2001]	32..645 37..650	591/614 (96%) 595/614 (96%)	0.0
AAB41584	Human ORFX ORF1348 polypeptide sequence SEQ ID NO:2696 - Homo sapiens, 663 aa. [WO200058473-A2, 05-OCT-2000]	1..664 1..663	377/671 (56%) 468/671 (69%)	0.0
ABB57299	Mouse ischaemic condition related protein sequence SEQ ID NO:839 - Mus musculus, 673 aa. [WO200188188-A2, 22-NOV-2001]	1..664 9..673	352/670 (52%) 457/670 (67%)	0.0
AAE06034	Mouse oocyte protein 5 (MOP5) - Mus musculus, 664 aa. [WO200153339-A2, 26-JUL-2001]	20..664 1..664	280/675 (41%) 402/675 (59%)	e-136

- 5 In a BLAST search of public sequence databases, the NOV16a protein was found to have homology to the proteins shown in the BLASTP data in Table 16D.

Table 16D. Public BLASTP Results for NOV16a				
Protein Accession Number	Protein/Organism/Length	NOV16a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9ULW8	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Homo sapiens (Human), 664 aa.	1..664 1..664	531/664 (79%) 565/664 (84%)	0.0
Q9Z184	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Mus musculus (Mouse), 664 aa.	1..664 1..664	484/664 (72%) 552/664 (82%)	0.0
P70708	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Rattus norvegicus (Rat), 664 aa.	1..664 1..664	481/664 (72%) 546/664 (81%)	0.0
O02849	Protein-arginine deiminase type III (EC 3.5.3.15) (Peptidylarginine deiminase III) - Ovis aries (Sheep), 664 aa.	1..664 1..664	485/664 (73%) 539/664 (81%)	0.0
Q9ULC6	Protein-arginine deiminase type I (EC 3.5.3.15) (Peptidylarginine deiminase I) - Homo sapiens (Human), 663 aa.	1..664 1..663	487/669 (72%) 544/669 (80%)	0.0

PFam analysis predicts that the NOV16a protein contains the domains shown in the Table 16E.

Table 16E. Domain Analysis of NOV16a			
Pfam Domain	NOV16a Match Region	Identities/ Similarities for the Matched Region	Expect Value
PAD	1..664	465/691 (67%) 623/691 (90%)	0

Example 17.

- 5 The NOV17 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 17A.

Table 17A. NOV17 Sequence Analysis			
	SEQ ID NO: 41	1027 bp	
NOV17a, CG172805-01 DNA Sequence	GTCGACGCGGGCCGCGGGCGCGGAGATGGCGGCGGGGATAGTGGCTTCTCGCAGACTCGCGACCTACTGACCCGGCGACTGACAGGCTCCAACCTACCCGGGACTCAGTATTAGCCTTCGCCTCACTGGCTCCTCTGCACAAGAGGCGGCTTCCGGAGTAGCCCTCGGTGAAGCCCAACACAGCTATGAGTCCCTTCGTGTGACGTCTGCGCAGAAACATGTTCTGCATGTCCAGCTCAACCGGCCAACAGAGGAATGCCATGAACAAGGTCTTCTGGAGAGAGATGGTAGAGTGCTTCAACAAGATTTTCGAGAGACGCTGACTGTCTGGGCGGTGGTGATCTCTGGTGCAGGAAAAATGTTCACTGCAGGTATTGACCTGATGGACATGGCTTCGGACATCTGCAGCCCCAAGGAGATGATGTGGCCCCGGATCAGCTGGTACCTCCGTGACATCATCATCTCGATACCAGGAGACCTTCAACGTCATCGAGAGGTGCCCCAAAGCCCGTGATTGCTGCGTCCATGGGGGTGCATTGGCGGAGGTGTGGACCTTGTCACCGCCTGTGACATCCGGTACTGTGCCCAGGATGCTTTCTTCCAGGTGAAGGAGGTGGACGTGGGTTTGGCTGCCATGTAGGAACACTGCAGCGCCTGCCCAAGGTCACTCGGGAACCAGAGCCCTGGTCAACGAGCTGGCCTTCACCGCCCGCAAGATGATGGCTGACGAGGCCCTGGGCAGTGGGCTGGTCAAGCCGGGTGTTCCAGACAAAGAGGTCACTGCTGGATGCTGCCTTAGCGCTGGCGGCCGAGATTCCAGCAAGAGCCCCGTGGCGGTGCAGAGCACCAAGGTCAACCTGCTGTATTCCCGGACCATTCGGTGGCCGAGAGCCTCAACTACGTGGCGTCTTGGAAACATGAGCATGCTGCAGACCAAGACCTCGTGAAGTCGGTCCAGGCCACGACTGAGAACAAGGAAGTGA		
	ORF Start: ATG at 26	ORF Stop: TGA at 1010	
	SEQ ID NO: 42	328 aa	MW at 35757.7kD
NOV17a, CG172805-01 Protein Sequence	MAAGIVASRRRLDLLTRRLTGSNYPGLSISLRLTGSSAQEAASGVALGEAPDHSYESIRVTSQAQKHVLHVQLNRPKNRNAMNKVFWREMVCEFNKISRADCRVVISGAGKMFTAGIDLMDMASDILQPKGDDVARISWYLRDIITRYQETFNVIERCPKPVIAAVHGGCIGGGVDLVTACDIRYCAQDAFFQVKEVDVGLAADVGTQLRQPKVIGNQSLVNLELAFTARKMMADEALGSGLVSRVFPDKEVMLDAALALAAEISSKSPVAVQSTKVNLLYSRDHSVAESLNYVASWNMSMLQTDLVKSVQATTENKELKTVTFSKL		

Further analysis of the NOV17a protein yielded the following properties shown in Table 17B.

Table 17B. Protein Sequence Properties NOV17a	
PSort analysis:	0.9200 probability located in mitochondrial matrix space; 0.8000 probability located in microbody (peroxisome); 0.6000 probability located in mitochondrial inner membrane; 0.6000 probability located in mitochondrial intermembrane space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV17a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 17C.

Table 17C. Geneseq Results for NOV17a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV17a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB43664	Human cancer associated protein sequence SEQ ID NO:1109 - Homo sapiens, 300 aa. [WO200055350-A1, 21-SEP-2000]	2..276 5..278	265/275 (96%) 268/275 (97%)	e-147
AAM24501	Colon tumour related predicted amino acid sequence for CT632 - Homo sapiens, 167 aa. [WO200149716-A2, 12-JUL-2001]	46..212 1..167	166/167 (99%) 166/167 (99%)	1e-93
AAB11904	Human colon tumour protein CT632, SEQ ID NO:204 - Homo sapiens, 167 aa. [WO200037643-A2, 29-JUN-2000]	46..212 1..167	166/167 (99%) 166/167 (99%)	1e-93
ABB61607	Drosophila melanogaster polypeptide SEQ ID NO 11613 - Drosophila melanogaster, 289 aa. [WO200171042-A2, 27-SEP-2001]	51..328 13..289	145/281 (51%) 199/281 (70%)	1e-73
AAG00231	Human secreted protein, SEQ ID NO: 4312 - Homo sapiens, 118 aa. [EP1033401-A2, 06-SEP-2000]	1..118 1..118	118/118 (100%) 118/118 (100%)	2e-61

In a BLAST search of public sequence databases, the NOV17a protein was found to have homology to the proteins shown in the BLASTP data in Table 17D.

Table 17D. Public BLASTP Results for NOV17a				
Protein Accession Number	Protein/Organism/Length	NOV17a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8WVX0	HYPOTHETICAL 35.8 KDA PROTEIN - Homo sapiens (Human), 328 aa.	1..328 1..328	328/328 (100%) 328/328 (100%)	0.0
Q96EZ9	UNKNOWN (PROTEIN FOR MGC:19561) - Homo sapiens (Human), 328 aa.	1..328 1..328	327/328 (99%) 327/328 (99%)	0.0
Q13011	Delta3,5-delta2,4-dienoyl-CoA isomerase, mitochondrial precursor (EC 5.3.3.-) - Homo sapiens (Human), 328 aa.	1..328 1..328	318/328 (96%) 320/328 (96%)	e-178
Q62651	Delta3,5-delta2,4-dienoyl-CoA isomerase, mitochondrial precursor (EC 5.3.3.-) - Rattus norvegicus (Rat), 327 aa.	1..328 1..327	248/328 (75%) 289/328 (87%)	e-143
A57626	peroxisomal enoyl hydratase-like protein - rat, 327 aa.	1..328 1..327	247/328 (75%) 288/328 (87%)	e-142

Pfam analysis predicts that the NOV17a protein contains the domains shown in the Table 17E.

Table 17E. Domain Analysis of NOV17a			
Pfam Domain	NOV17a Match Region	Identities/ Similarities for the Matched Region	Expect Value
ECH	68..248	66/185 (36%) 132/185 (71%)	1e-41

Example 18.

- 5 The NOV18 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 18A.

Table 18A. NOV18 Sequence Analysis			
	SEQ ID NO: 43	1067 bp	
NOV18a, CG56763-01 DNA Sequence	CCCTTTTCCTCTTGCTCTTTGATGTTTTGTAGGCCTGCAGCTCCCAAGCACAGAGGCAT GAGTGGGGAGAATGTCAACAGGGTCGGCACCTTCATCTGGTGGGCTTCCCCACGGCC CCAGGGCTGCAGTACCTGCTCTTCTCTCTCTTCTCTGCTCACCTACCTCTTTGTCTCTGG TGGAGAACCTGGCCATCATCCTCACCGTCTGGAGCAGCACCTCCCTCCACAGGCCCAT GTACTACTTTCTGAGCTCCATGTCTTTCTAGAGATCTGGTACGTGTCTGACATCACC CCCAAGATGCTGGAGGGCTTCTCTCTCCAGCAGAAACGCATCTCTTTCTGCTCGGGTGCA TGACGCAGCTCTACTTCTTCTCAGCTCCCTGGTGTGCACCGAGTGTGTGCTTCTGGCCTC CATGGCCTACGACCGCTACGTGGCCATCTGCCACCCGCTGCGCTACCACGCTCCTTGTG ACCCCGGGGCTGTGCCTCCAGCTGGTGGGCTTCTCCTTTGTGAGTGGCTTACCATCT CCATGATCAAGGTCTGTTTTATCTCCAGCGTCACGTTCTGTGGCTCCAACGTCTTGAA CCACTTCTTCTGTGACATTCCCCCATCCTCAAGCTGGCCTGCACGGACTTCTCCACT GCAGAGCTGGTGGATTTCATTCTGGCCTTCATCATCTGGTGTTCCTCACTCCTGGCCA CCATGCTGTATATGCGCACATCACCCTGGCTGTCTGCGCATCCCTCGGCCACCGG CTGCTGGAGAGCCTTCTTACCTGCGCCTCTCACCTACCGTGGTCAACGCTCTTCTAT ACAGCCTTGCTTTTTCATGTATGTCCGGCCCCAGGCCATTGATTCCCGAGCTCCAACA AGCTCATCTCTGTTTTGTACACAGTTATACCCCATCTTGAACCCCTTGATATACTG CCTGAGGAATAAGGAATTTAAGAAATGCCTTGAAAAAGCCTTCGGCTTGACGAGCTGC GCCGTAGAGGGAGGCTTCTAGTCTTCTGGAACCTCATCTCCAATACACAGCCAGC CTCTCTGAGGAGGCCATTGACT		
	ORF Start: ATG at 21	ORF Stop: TGA at 1050	
	SEQ ID NO: 44	343 aa	MW at 38584.5kD
NOV18a, CG56763-01 Protein Sequence	MFCRPAAPKHRGMSGENVTRVGTFILVGFPPTAPGLQYLLFLLFLLTYLFVLVENLAI LTWVSSTSLHRPMYFLLSMSFLEIYWVSDITPKMLEGFLQKKRISFVGCMTQLYFF SSLVCTECVLLASMAYDRYVAICHPLRYHVLVTPGLCLQLVGFVSFVSFTISMIVCF ISSVTFCGSNVLNHFFCDISPILKLACTDFSTAELVDFILAFIILVFPPLATMLSYAH ITLAVLRIPSATGCWRAFFTCAHLTVTVTFYTALEFMYVRPQAIDSRSSNKLSVLY TVITPILNPLIYCLRNKEFKNALKKAFGLTSCAVEGRLSSLLEHLQIHSQPL		

Further analysis of the NOV18a protein yielded the following properties shown in Table 18B.

Table 18B. Protein Sequence Properties NOV18a	
PSort analysis:	0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.3700 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 64 and 65

A search of the NOV18a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 18C.

Table 18C. Geneseq Results for NOV18a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV18a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG71967	Human olfactory receptor polypeptide, SEQ ID NO: 1648 - Homo sapiens, 319 aa. [WO200127158-A2, 19-APR-2001]	13..332 1..319	304/320 (95%) 308/320 (96%)	e-173
AAG71962	Human olfactory receptor polypeptide, SEQ ID NO: 1643 - Homo sapiens, 314 aa. [WO200127158-A2, 19-APR-2001]	13..319 1..307	289/307 (94%) 299/307 (97%)	e-167
ABG16907	Novel human diagnostic protein #16898 - Homo sapiens, 291 aa. [WO200175067-A2, 11-OCT-2001]	5..295 20..289	267/291 (91%) 267/291 (91%)	e-150
ABG16907	Novel human diagnostic protein #16898 - Homo sapiens, 291 aa. [WO200175067-A2, 11-OCT-2001]	5..295 20..289	267/291 (91%) 267/291 (91%)	e-150
AAG71675	Human olfactory receptor polypeptide, SEQ ID NO: 1356 - Homo sapiens, 327 aa. [WO200127158-A2, 19-APR-2001]	15..313 6..308	177/303 (58%) 224/303 (73%)	e-100

In a BLAST search of public sequence databases, the NOV18a protein was found to have homology to the proteins shown in the BLASTP data in Table 18D.

Table 18D. Public BLASTP Results for NOV18a				
Protein Accession Number	Protein/Organism/Length	NOV18a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8VGU5	OLFACTORY RECEPTOR MOR103-2 - Mus musculus (Mouse), 312 aa.	13..319 1..307	270/307 (87%) 287/307 (92%)	e-156
Q8VGU4	OLFACTORY RECEPTOR MOR103-3 - Mus musculus (Mouse), 312 aa.	13..319 1..307	261/307 (85%) 284/307 (92%)	e-153
Q9EPG2	M51 OLFACTORY RECEPTOR - Mus musculus (Mouse), 314 aa.	17..322 5..310	211/306 (68%) 250/306 (80%)	e-124
Q8VGU1	OLFACTORY RECEPTOR MOR103-4 - Mus musculus (Mouse), 314 aa.	17..322 5..310	210/306 (68%) 249/306 (80%)	e-124
Q9EPV0	M50 OLFACTORY RECEPTOR (OLFACTORY RECEPTOR M50) (OLFACTORY RECEPTOR MOR103-16) - Mus musculus (Mouse), 316 aa.	16..315 2..301	209/300 (69%) 246/300 (81%)	e-121

PFam analysis predicts that the NOV18a protein contains the domains shown in the Table 18E.

Table 18E. Domain Analysis of NOV18a			
Pfam Domain	NOV18a Match Region	Identities/ Similarities for the Matched Region	Expect Value
7tm_1	53..302	50/269 (19%) 181/269 (67%)	6.5e-27
Colicin_V	147..310	29/197 (15%) 100/197 (51%)	0.43

Example 19.

- 5 The NOV19 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 19A.

Table 19A. NOV19 Sequence Analysis

Table 19A. NOV19 Sequence Analysis			
	SEQ ID NO: 45	1071 bp	
NOV19a, CG56777-01 DNA Sequence	GAAAAGGAGGAAAGACAGGTTAGCTGGGGATGATGACGGATCTGAAGCAAAGCCATT AGTGAGGCTGAATGATGGACCCCTCATGCCAGTGCTGGGATTGGCACTTATGCTCC GATCATACTCCAAAGCCAGGCTGCCGAGGCCACCAAAGTGGCTATTGACGTAGGC TCCGCCATATTGATTGATCAGCATACTTATACCAAATGAGGAGGAGTTGGACAGGCC TTGGGAGAAGATCGCTGATGGTACCGTCAAGAGAGAGAAATATTCTACACCATCAAG CTTTGGGCTACTTTCTTTTCGGGCAGAATTGGTTACCCGGCCCTAGAAAGGTCAC AGAAACTTGGACCGGACTATGTAGATCTCTTCATTATTCATGTACCATTTGCTATGA GCCTGGGAAAGAATTACTGCCAAAGGATGCCAGTGGAGAGATTATTTAGAACTGTG GAGCTTTGTGACACTTGGGAGGCCCTGGAGAAGTGCAAAGAGCAGGTTTAAACAGGT CCATTGGGGTGTCCAATTTCAATCACAAGCTGCTGGAAGTCTATCTCAACAAGCCAG GCTCAAGTACAAGCCCACTGCAACCAGGTGGAATGTCAACCCTTACCTCAACCAAGAC AAACTCCTGGAGTTCTGCAAGTCCAAGGACATTGTTCTAGTTGCTACAGTGCCCTGC GATCCCAAAGAGACCCACAGTGGGTGGATCCCGACTGCCACATCTCTTGGAGGAGCC GATCTTGAAATCCATTGCCAAGAAACACAGTGGAAAGCCAGCCAGGTGCGCCTGCGC TACCAGCTGCAGCGGGAGTGGTGGTGTCTGGCCAAGAGCTTCTCTCAGGAGAGAATCA AAGAGAATCTCCAGATTTTGAAGTTGAGTTGACTCCAGAGGACATGAAAGCCATTGA TGGCCTCAACAGAAATCTCCGATATGACAGTTACAATTTGCTGCTAATCACCCCTTAT TTTCCATTTTCTGAAGAATATTGACCATGAGCTATTGAACATTACCACCACAATTTCC TCCTTTTAGGTGACAGAGAGAGGATT		
	ORF Start: ATG at 30	ORF Stop: TGA at 1008	
	SEQ ID NO: 46	326 aa	MW at 37132.2kD
NOV19a, CG56777-01 Protein Sequence	MMTDLKQSHSVRLNDGPFMPVLGFGTYAPDHTPKSQAAEATKVAIDVGFRIHDSAYLY QNEEEVGQAIWEKIADGTVKREEIFYTIKLWATFFRAELVHPALERSLKKLGPDYVDL FIIHVPFAMKPGKELLPKDASGEIILETVELCDTWEALEKCKEAGLTRSIGVSNFNH LLELILNKPGLKYKPTCNQVECHPYLNQSKLLEFCKSKDIVLVAYSALGSQRPQWVDL PDCPHLLEEPILKSIKKHSGSPGQVALRYQLQRGVVVLAKSFQERIKENFQIPDFE LTPEDMKAIDGLNRNLRYDKLQFAANHPYFPFSEY		
	SEQ ID NO: 47	1108 bp	
NOV19b, CG56777-02 DNA Sequence	GAAAAGGAGGAAAGACAGGTTAGCTGGGGATGATGACGGATCTGAAGCAAAGCCATT AGTGAGGCTGAATGATGGACCCCTCATGCCAGTGCTGGGATTGGCACTTATGCTCCT GATCATACTCCAAAGCCAGGCTGCCGAGGCCACCAAAGTGGCTATTGACGTAGGCT TCCACCATATTGATTGATCAGCATACTTATACCAAATGAGGAGGAGTTGGACAGGCCAT TTGGGAGAAGATCGCTGATGGTACCGTCAAGAGAGAGGAAATATTCTACACCATCAAG CTTTGGGCTACTTTCTTTTCGGGCAGAATTGGTTACCCGGCCCTAGAAAGGTCAC AGAAACTTGGACCGAGACTATGTAGATCTCTTCATTATTCATGTACCATTTGCTATGA GGTCAGGTTCTTCTGAGAGAGGCAGATTGAGTGAGGTGCATAGGATTGCCAAACT TTCCCTGGGAAAGAATTACTGCCAAAGGATGCCAGTGGAGAGATTATTTAGAACTG TGGAGCTTTGTGACACTTGGGAGGCCCTGGAGAAGTGCAAAGAAGCAGGTTTAAACAG GTCCATTGGGGTGTCCAATTTCAATCACAAGCTGCTGGAAGTCTATCTCAACAAGCCA GGGCTCAAGTACAAGCCCACTGCAACCAGGTGGAATGTCAACCCTTACCTCAACAGCA GCAAACTCCTGGAGTTCTTCAAGTCCAAGGACATTGTTCTAGTTGCCTACAGTGCCCT GGGATCCCAAAGAGACCCACAGTGGGTGGATCCCGACTGCCACATCTCTTGGAGGAG CCGATCTTGAAATCCATTGCCAAGAAACACAGTGAAGCCCAAGGCCAGGTGCGCCTGC GCTACCAGCTGCAGCGGGAGTGGTGGTGTCTGGCCAAGAGCTTCTCTCAGGAGAGAAT CAAAGAGAACTTCCAGATTTTGAAGTTGAGTTGACTCCAGAGGACATGAAAGCCATT GATGGCCTCAACAGAAATCTCCGATATGATAAGTTACAATTTGCTGCTAATCACCCCT ATTTTCCATTTTCTGAAGAATATTGATCACCCCTTATTTTCCATTTTCTGAAGAATATT GATCAC		
	ORF Start: ATG at 30	ORF Stop: TGA at 1068	
	SEQ ID NO: 48	346 aa	MW at 39673.1kD
NOV19b, CG56777-02 Protein Sequence	MMTDLKQSHSVRLNDGPFMPVLGFGTYAPDHTPKSQAAEATKVAIDVGFHHIDSAYLY QNEEEVGQAIWEKIADGTVKREEIFYTIKLWATFFRAELVHPALERSLKKLGPDYVDL FIIHVPFAMKVRFFLREADLSEVHRIAKTFPGKELLPKDASGEIILETVELCDTWEAL EKCKEAGLTRSIGVSNFNHKLLELILNKPGLKYKPTCNQVECHPYLNQSKLLEFFKSK DIVLVAYSALGSQRPQWVDPDCPHLLEEPILKSIKKHRSRPGQVALRYQLQRGVV LAKSFQERIKENFQIPDFELTPEDMKAIDGLNRNLRYDKLQFAANHPYFPFSEY		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 19B.

Table 19B. Comparison of NOV19a against NOV19b.		
Protein Sequence	NOV19a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV19b	1..326 1..346	323/346 (93%) 323/346 (93%)

Further analysis of the NOV19a protein yielded the following properties shown in Table 19C.

Table 19C. Protein Sequence Properties NOV19a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.4275 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV19a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 19D.

Table 19D. Geneseq Results for NOV19a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV19a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM79455	Human protein SEQ ID NO 3101 - Homo sapiens, 325 aa. [WO200157190-A2, 09-AUG-2001]	4..326 4..325	220/323 (68%) 269/323 (83%)	e-131
AAM78471	Human protein SEQ ID NO 1133 - Homo sapiens, 323 aa. [WO200157190-A2, 09-AUG-2001]	4..326 2..323	220/323 (68%) 269/323 (83%)	e-131
AAW14799	Type 5 17-beta-hydroxysteroid dehydrogenase - Homo sapiens, 323 aa. [WO9711162-A1, 27-MAR-1997]	4..326 2..323	220/323 (68%) 269/323 (83%)	e-131
AAB43444	Human cancer associated protein sequence SEQ ID NO:889 - Homo sapiens, 336 aa. [WO200055350-A1, 21-SEP-2000]	1..326 10..336	216/327 (66%) 268/327 (81%)	e-128
AAB76865	Human lung tumour protein related protein sequence SEQ ID NO:783 - Homo sapiens, 364 aa. [WO200100828-A2, 04-JAN-2001]	33..326 71..364	200/294 (68%) 248/294 (84%)	e-120

In a BLAST search of public sequence databases, the NOV19a protein was found to have homology to the proteins shown in the BLASTP data in Table 19E.

Table 19E. Public BLASTP Results for NOV19a				
Protein Accession Number	Protein/Organism/Length	NOV19a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P05980	Prostaglandin-F synthase 1 (EC 1.1.1.188) (PGF synthase 1) (PGF 1) (Prostaglandin-D2 11 reductase 1) (PGFSI) - Bos taurus (Bovine), 323 aa.	7..326 4..323	229/320 (71%) 270/320 (83%)	e-134
P52897	Prostaglandin-F synthase 2 (EC 1.1.1.188) (PGF synthase 2) (PGF 2) (Prostaglandin-D2 11 reductase 2) (PGFSII) - Bos taurus (Bovine), 323 aa.	7..326 4..323	228/320 (71%) 268/320 (83%)	e-134
P52898	Dihydrodiol dehydrogenase 3 (EC 1.-.-.-) (Prostaglandin F synthase) - Bos taurus (Bovine), 323 aa.	11..326 8..323	226/316 (71%) 264/316 (83%)	e-133
Q9NS54	3ALPHA-HYDROXYSTERIOD DEHYDROGENASE VARIANT (EC 1.1.1.213) - Homo sapiens (Human), 323 aa.	11..326 8..323	216/316 (68%) 265/316 (83%)	e-131
P80508	Prostaglandin-E2 9-reductase (EC 1.1.1.189) (20-alpha-hydroxysteroid dehydrogenase) (EC 1.1.1.149) (20-alpha-HSD) - Oryctolagus cuniculus (Rabbit), 323 aa.	11..326 8..323	219/316 (69%) 259/316 (81%)	e-131

Pfam analysis predicts that the NOV19a protein contains the domains shown in the Table 19F.

Table 19F. Domain Analysis of NOV19a			
Pfam Domain	NOV19a Match Region	Identities/ Similarities for the Matched Region	Expect Value
aldo_ket_red	13..306	164/368 (45%) 261/368 (71%)	1.5e-144

Example 20.

- 5 The NOV20 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 20A.

Table 20A. NOV20 Sequence Analysis			
	SEQ ID NO: 49	979 bp	
NOV20a, CG56941-01 DNA Sequence	TGGAAAACCTGTGCAGTTTAGGACAACCTCTCTTTGTAAAATTAGTGGTTCTTGGGAG TCAAAATGCTTCCTGGTGCCGGGAGTGAGCGATGAGCCGGCTTCTGTCCCTGGCGGC AGAGTCGCCTTGGCCGCCTTGCCTGCCCGCGCTCTCGCGGGTTCGGGATGTTCTATG CCGTGAGGAGGGGCGGCAAGTCCGCGGTCTTTCTGACCGGGAATGAGTGCAAAGCACA GGTGGACCGGTTTCTGCTGCCAGATTAAAGAAGTTTGCCACAGAGGATGAGGCCTG GACTTTGTGAGGAAATCTGCAAGCCCGGAAGTTTCAGAAGGGCAGGAAATCAACATG GACAAGAATCGGAGACGAAAGCCAGCAAGCGACTCCGTGAGCCACTGGATGGAGATG AGATGAAAGCGCAGAGCCATATGCAAGCACAATGAAGCCGAGCGTGAAGCCGGCGCT CCAGTTAGCAGAGACACGTTTCTTACATGGGAGACTTCGTCGTCGTCTACGCTGATG GCTGCTGCTCCAGTAATGGGCGTAGAAAGCCACGAGCAGGAATCCGCGTTTACTGGGG GCCGGGCTATCCTTTAACTGTAGGCATTAGACTTCCTGGGTGGCAGACAAACCAAAGA GCGGAAATTCATGCAGCCTGCAAGGCCATTGAACAAGCACAGACTCAGAAAATCAAT AACTGGTTCTGTATACAGACAGTATGTTTACGATAAATGGTATACTAACTGGGTTCT AGGTTGGAAGAAAAATGGGTGGAAGACAAGTGCAGGGAAGAGGTGATCAACAAGAG GACTTCGTGGCACTGGAGAGGCTGACCCAGGGGATGGACATTTCAGTGGGCGAGTATC TAAATGTTTCATGTCCCATGGCTATTGTGACCAAATAAGCTAAGTTTACTAGAGATTAA TGGTTTATGTAAGTGGTTTAATGACATAACAATTTCCATGAAGTTAATATC		
	ORF Start: ATG at 90	ORF Stop: TGA at 897	
	SEQ ID NO: 50	269 aa	MW at 29980.9kD
NOV20a, CG56941-01 Protein Sequence	MSRLLSLAARVALAALPCRRSRGFGMFYAVRRGGKSAVFLTGNECKAQVDRFPAARF KFATEDEAWDFVRKSASPEVSEGEQHQEQSETKASKRLREPLDGDGDESAEPPYAS MKPSVKPAPPVSRDTFSYMGDFVVVYADGCCSSNGRRKPRAGIRVYWGPGYPLTVGIR LPGWQTNQRAEIIHAACKAIEQAQTQKINKLVLYTDSMFTINGITNWWQGWKKNGWKT AGKEVINKEDFVALERLTQGMIDIQWASILNVHVPWLL		

Further analysis of the NOV20a protein yielded the following properties shown in Table 20B.

Table 20B. Protein Sequence Properties NOV20a	
PSort analysis:	0.3700 probability located in outside; 0.2339 probability located in microbody (peroxisome); 0.1080 probability located in nucleus; 0.1000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 24 and 25

A search of the NOV20a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 20C.

Table 20C. Geneseq Results for NOV20a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV20a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY70235	Human RNA-associated protein-16 (RNAAP-16) - Homo sapiens, 286 aa. [WO200011171-A2, 02-MAR-2000]	1..266 1..262	236/267 (88%) 244/267 (90%)	e-134
AAB97508	Human type II RNase H protein - Homo sapiens, 286 aa. [WO200123613-A1, 05-APR-2001]	1..266 1..262	235/267 (88%) 242/267 (90%)	e-134
AAY25094	Human type 2 RNase H protein - Homo sapiens, 286 aa. [WO9928447-A1, 10-JUN-1999]	1..266 1..262	235/267 (88%) 242/267 (90%)	e-134
AAB97511	E coli type II RNase H protein - Escherichia coli, 216 aa. [WO200123613-A1, 05-APR-2001]	23..266 4..206	157/244 (64%) 178/244 (72%)	2e-85
AAY25098	Mouse RNase H homologue protein fragment - Mus sp, 216 aa. [WO9928447-A1, 10-JUN-1999]	23..266 4..206	157/244 (64%) 178/244 (72%)	2e-85

In a BLAST search of public sequence databases, the NOV20a protein was found to have homology to the proteins shown in the BLASTP data in Table 20D.

Table 20D. Public BLASTP Results for NOV20a

Protein Accession Number	Protein/Organism/Length	NOV20a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
O60930	Ribonuclease H1 (EC 3.1.26.4) (RNase H1) (Ribonuclease H type II) - Homo sapiens (Human), 286 aa.	1..266 1..262	237/267 (88%) 244/267 (90%)	e-134
Q8VCR6	RIBONUCLEASE H1 - Mus musculus (Mouse), 285 aa.	1..266 1..261	183/267 (68%) 213/267 (79%)	e-101
O70338	Ribonuclease H1 (EC 3.1.26.4) (RNase H1) - Mus musculus (Mouse), 285 aa.	1..266 1..261	183/267 (68%) 213/267 (79%)	e-101
Q91953	MRNA, COMPLETE CDS, CLONE CLFEST65 - Gallus gallus (Chicken), 293 aa.	11..266 6..269	141/270 (52%) 178/270 (65%)	2e-71
Q21024	F59A6.6 PROTEIN - Caenorhabditis elegans, 369 aa.	27..260 67..342	76/276 (27%) 122/276 (43%)	2e-25

PFam analysis predicts that the NOV20a protein contains the domains shown in the Table 20E.

Table 20E. Domain Analysis of NOV20a			
Pfam Domain	NOV20a Match Region	Identities/ Similarities for the Matched Region	Expect Value
RnaseH	136..257	50/147 (34%) 93/147 (63%)	1.1e-27

Example 21.

The NOV21 clone was analyzed, and the nucleotide and encoded polypeptide
5 sequences are shown in Table 21A.

Table 21A. NOV21 Sequence Analysis			
	SEQ ID NO: 51	2536 bp	
NOV21a, CG57109-01 DNA Sequence	GGGACACTGACATGGACTGAAGGAGTAGAAAAAGAAGCCTTGGGCTCTCCAGATGGAA GAATGACCGTGTGAGGAAACTGTTTAACTCAAGGGCAGGGAAATCAGGAGCGTCTCT GATTTCTTCAGGGAAGGGGATGCTTTTCATAGCTATGGGCAAAGAACCTGACACTGA AGAGCATTGAGGTGGCTGTAGAAGAACTGTACCCCAACAAAGCCCGGCCCTGACACT GGCCAGCACAGCCGTGCCCTTCTCCAAGGCTGAGGAGCAGGCTGTTTAGCAAGGCT CTGAAAGGAGACCACCGCTGTGGGGAGACCCGAGACCCCAAGAGCTGCAGCGAAGTTG CAGGATGCAAGGCAGCCATGAGGCACCAGGGGAAGATCCCCGAGGAGCTTTCAC TAGA TGACAGAGCGAGGACCCAGAAGAAGTGGGGGAGGGGAAATGGGAGCCAGAACCCAGT AGCAAGCCCCCAGGGAAGCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAGAGCATC TTGGGGTGGAGATTGAAAAGACCTCGGGTGAAATTATCAGATGCGAGAAGTGCAAGAG AGAGAGGGAGCTTCAGCAGAGCCTGGAGCGTGAGAGGCTTCTCTGGGGACCAAGTGA CTGGATATGGGGAAGGGCCCAATGTATGATGTGGAGAAGCTGGTGAGGACCAAGCT GCAGGAGGTCTCCCGAGGCAAATCCTGCAAGTGGGGAGGAAGGTTGGAAGGGTGACAG CCACAGGAGCAGCCCCAGGAATCCCACTCAAGAGCTGAGGAGACCCAGCAAGAGCATG GACAAGAAAGGACAGAGGCCAGAGGATCAAGAAAGCCATGTCTCAGGGAGCAGCCA AGGCCAAGAAGGACCTTGTGGAAGTTCTTCTGTCTCACAGAGGAGGGGCTGAGGGAGGT GAAGAAGGACACCAGGCCCATGAGCAGGAGCAAAACATGGTGGCTGGCTCTCTGAGAGAG CACCAGGCGGGCTTTGAGAAGCTCCGAGGACCCGAGGAGAAGAGAAGGAGGCAGAGAG AGGAGAAAAAGCCATGTATGTCTGGAGGCAGAAGGATGACTCTCAGAGATGACCAACC TGCAAGCTAGAAAAGGAGCCCAAGACGAGGCCAGAAGAGAACAGCCAGAGCGGGCCC AGCGGTTCGGAAGCCACGGCCCATGGGCATCATTGCCGCCAATGTGGAAGAGCATTATG AGACTGGCCCGGGTCAATGGGGATGGGAACCTTGTCTGTCGTGAAGGAGTGACAGACCCG CGAGACCAGGCAGGCCCTATGCGATGAAGATCATTGACAAGTCCAGACTCAAGGGCAAG GAGGACATGGTGGACAGTGAGATCTTGATCATCCAGAGCCTCTCTACCCCAACATCG TGAAATTGTCATGAAGTCTACGAAACAGACATGGAATCTACCTGATCCTGGAGTACGT GCAGGGAGGAGACCTTTTTCAGCCCATCATAGAAAGTGTGAAGTCCCGGAGCCCGAT GCTGCCCTCATGATCATGGACTTATGCAAAAGCCCTCGTCCACATGCACGACAAGAGCA TTGTCCACCGGGACCTCAAGCCGGAAAAACCTTTTGGTTTCAGCGAAATGAGGACAAATC TACTACTTGAAATTGGCTGATTTTGGACTTGCAAGCATGTGGTGAGACCTATATTT ACTGTGTGTGGGACCCCACTTACGTAGCTCCCGAAATCTTTCTGAGAAAGGTTATG GACTGGAGGTGGACATGTGGGCGGCTGGCGTGATCCTCTATATCCTGTGTGTGGCTT TCCCCATTCCGAGCCCTGAGAGGGACACGAGCAGCTCTTAACATCATCCAGCTG GGCCACTTTGAGTTTCTCCCCCTTACTGGGACAATATCTCTGATGCTGCTAAGATC TGGTGAGCCGGTTGCTGGTGGTAGACCCAAAAAGCGCTACACAGCTCATCAGGTTCT TCAGACCCCTCGGATCGAAACAGCTGGCAAGACCAATACAGTGAACGACAGAGAAGCAG GTGTCCCCCAGCAGCGAGGGTCACCTCCGGAGCCAGCACAAGAGGGTTGTGGAGCAGG TATCATAGTACCACCTTGGGAATCTGTCCAGCCCCAGTCTGTCTCAGGACAGAGAGA AAAGGATAGAAAGTTTGAGAGAAAAACAATGAAAGAGGCTTCTTCACATAATTGGTGAA TCAGAGGGAGAGACACTGAGTATATTTTAAAGCATATTAATAAATTAAGTCAATGTT AAATGTCACAACATATTTTAGATTGTATATTTAAAGCCTTTAATACATTTTGGGG GGTAAGCATTGTCTCAGTGAGGAATTTTGGTAATAATGATGTGTTTGTCTTCCCTT TGTAACCAAGTTTATTCTGTACTACAGGAGTGGTGTCTTACAGGGGTCTAAACTCCCC TGTGAGATTAAATAAGGTGCATTGTGCTTTCTGTGTTAATAAATGTGCTCTGAATA ACAGAAAAAAGG		
	ORF Start: ATG at 150	ORF Stop: TAG at 2094	
	SEQ ID NO: 52	648 aa	MW at 73813.6kD
NOV21a, CG57109-01 Protein Sequence	MGKEPLTLKSIQVAVEELYPNKARALTQAHSRAPSPRLRSRLFSKALKGDHRCGETE TPKSCSEVAGCKAAMRHQGGKIPEELSLDDRARTQKKWGRGWEPSPSSKPPREATLEE RHARGEKHLGVEIEKTSGEIIRCEKCKRERELQQLERERLSLGTSELDMGKGPMDV EKLVRTRSCRSPPEANPASGEEGWKGDSSHRSSPRNPTQELRRPSKSMKKEDRGPEQ ESHAQGAAKAKDLVEVLPVTEEGLREVKKDTRPMSRSKHGGWLLREHQAGFEKLRR RGEEKEAEKEKPKCMGGRRMTLRDDQPAKLEKEPKTRPENKPERPSGRKPRPMGII AANVEKHYETGRVIGDGNFAVVKECRHRETRQAYAMKIIDKSLKKGKEDMVDSEILII QSLSHPNIVKLHEVYETDMEIYLILEYVQGGDLFDALIESVKFPEPDAALMIMDLCKA LVHMHDKSIVHRDLKPNLLVQRNEDKSTTLKLADFLAKHVVRPIFTVCGTPTYVAP EILSEKGYGLEVDMWAAGVILYILLCGFPFPRSPERDQDELFINIQLGHFELFPPYVD NISDAAKDLVSRLLVDPKKRYTAHQVLQHPWIEAGKTNVTKRQKQVSPSSSEGHFRS QHKRVVEQVS		
	SEQ ID NO: 53	2808 bp	
NOV21b, CG57109-02 DNA Sequence	TTTACAGAGTCAGGCCTCACCCTGAGAGGGCTCCTGTATTAGTCCCTTTTCATGCTGC TTATAGAGACATACCTGAGACTGGGCAATTTCAGAGAAAGGTTTCTTGGACTTACA GTTAGTTCACAGTGGCTGGGAAGCCTCACAATCATGGCGGAAGGCAAGGAAGGGCAA GTCCCATCTTACATGGATGGCAGCAGGCAAGAGAGAATGAGGAAGATGCAAAAGCGG		

	AAACCCCTGATGTAACCATCAGATCTTATGAGATTTATTCACTACCATGGAACAGACA GCAAGGCGTATGTGACCATTTCTCTAGAATATTTAAGCTCGAGAATCTCAGAGCGGAAG CTGCAAGGCTCTCTGGCTGCTGCGAGCCGAGGGAATCTGGAGAAACCATTCCTGGGGC CGCGTGGCCCCGTCGTGCCCTTGTCTGCCCTCGGAATGGCCTTCACTCAGCACATCC TGAGAACAGCCCTCTGAAGCCAGGGTCTGACCGTAGTGAAGCTGGGTGGGCAGCGC CCCCGAAAGATCACTCTGCTCCTCAACAGGCGATCAGTGCAGACGTTTCAGAGCAGCTCT TAGCTGACATCTCAGAAGCCTTGGGCTCTCCAGATGGAAGAATGACCGTGTGAGGAA ACTGTTTAACTCAAGGCGAGGAAATCAGGAGCGTCTCTGATTTCTTCAGGGAAGGG GATGCTTTTATAGCTATGGGCAAGAACCCTGACACTGAAGAGCATTCAGGTGGCTG TAGAAGAACTGTACCCCAACAAAGCCCGGGCCTGACACTGGCCAGCAGACCGCTGC CCCTTCTCCAAGGCTGAGGAGCAGGCTGTTTAGCAAGGCTCTGAAGGAGACCAACCGC TGTGGGGAGACCGAGACCCCAAGAGCTGCAGCGAAGTTGAGGATGCAAGGCAGCCA TGAGGCACCAGGGGAAGATCCCCGAGGAGCTTTACTAGATGACAGAGCGAGGACCCA GAAGAAGTGGGGGAGGGGGAATGGGAGCCAGAACCCAGTAGCATGCCCCAGGGAA GCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAAGCATTTGGGGTGGAGATTGAAA AGACCTCGGGTGAATTTATCAGATGCGAGAAGTGCAGAGAGAGAGGGAGCTCCAGCA GAGCCTGGAGCGTGAGAGGCTTTCTCTGGGGACCAAGTGAAGTGGATATGGGGAAGGGC CCAATGTATGATGTGGAGAAGCTGGTGAGGACAGAACTGTCAGGAGGCTCTCCGAGG CAATCCTGCAAGTGGGGAGGAAAGGGTGAAGGGTGACGCCACAGGAGCAGCCCCAG GAATCCCACTCAAGAGCTGAGGAGACCCAGCAAGAGCATGGACAAGAAAGAGGACAGA GGCCAGAGGATCAAGAAAGCCATGCTCAGGGAGCAGCCAGGCCAAGAGGACCTTG TGGAAGTTCTTCTGTACAGAGGAGGGGCTGAGGGAGGTGAAGAAGGACACCGGCC CATGAGCAGGAGCAAAATGGTGGCTGGCTCCTGAGAGAGCACCAGCGGGCTTTGAG AAGCTCCGAGGACCCGAGGAGAAGAGAAGGAGGCAGAGAAGGAGAAAAAGCCATGTA TGTCTGGAGGCAGAAAGATGACTCTCAGAGATGACCAACCTGCAAGCTAGAAAAGGA GCCAAGACGAGGCCAGAAGAGAACAAGCCAGAGCGGCCAGCGGTGGAAGCCACGG CCCATGGGCATCATTTGCCGCAATGTGGAAGCATTATGAGACTGGCCGGGTCTATTG GGGATGGGAACCTTGCTGTCTGTAAGGAGTGACAGACCCGAGACCAAGGCAGGCCTA TGCGATGAAGATCATTGACAAGTCCAGACTCAAGGGCAAGGAGGACATGGTGAGCAGT GAGATCTTGATCTCCAGAGCCTCTCTCACCCCAACATCGTGAATTTGATGAAGTCT ACGAAACAGACATGGAATCTACCTGATCCTGGAGTACGTGCAGGGAGGAGACCTTTT TGACGCCATCATAGAAAGTGTGAAGTTCCTCGAGCCGATGCTGCCCTCATGATCATG GACTTATGCAAAGCCCTCGTCCACATGCACGACAAGAGCATTGTCCACCGGACCTCA AGCCGGAACCTTTTGGTTTCCGGAATGAGGACAAATCTACTACCTTGAAATGGC TGATTTTGGACTTGCAAAGCATGTGGTGAGACCTATTTACTGTGTGTGGGACCCCA ACTTACGTAGCTCCGAAATCTTTCTGAGAAAGGTTATGGACTGGAGGTGGACATG GGGTGCTGGCGTGATCTCTATATCCTGCTGTGTGGCTTTCCCTTCCGAGGCC TGAGAGGGACAGGACGAGCTCTTTAACATCATCCAGCTGGGCCACTTTGAGTTCCTC CCCCCTTACTGGGACAATATCTCTGATACAGCTGCTAAAGATCTGGTGAGCCGGTGC TGGTGGTAGACCCAAAAAGCGCTACACAGCTCATCAGGTTCTTCAGCACCCCTGGAT CGAAACAGCTGGCAAGCAATACAGTGAAACGACAGAAAGCAGGTGTCCTCCAGCAGC GAGGGTCACTTCCGAGCCAGCACAGAGGGTTGTGGAGCAGGTATCATAGTACCAC CTTGGGAATCTGTCCAGCCCCAGTCTGCTCAAGGACAGAGAAAGGATAGAAAGTTT GAGAGAAAAACAATGAAGAGGCTTCTTACATAATTGGTGAATCAGAGGGAGAGACA CTGAGTATATTTTAAAGCATATTA
	ORF Start: ATG at 151
	ORF Stop: TAG at 2659
	SEQ ID NO: 54
	836 aa
	MW at 95152.6kD
NOV21b, CG57109-02 Protein Sequence	MAEGKEGVPSYMDGSRQRENEEDAKAETPDVTIRSYEYISLPWNRQQGVCDHSLEYL SSRISERKLQGSWLPASRGNLEKPFLLGRGPVVPFLPCPRNGLHSAHPENSPLKPRVVT VVKLGGQRPRKITLLNRRSVQTFEQLLADI SEALGSPRWKNDVRKLFNLKGREIRS VSDFFREGDAFIAMGKEPLTLKSIQVAVEELYPNKARALTLAQHSRAPSPRLRSRLFS KALKGDHRCGETETPKSCSEVAGCKAAMRHQGKIPEELSLDDRARTQKKWGRGKWEPE PSSKPPREATLEERHARGEKHLGVEIEKTSGEIIRCEKCRERELQSLERERLSLGT SELDMGKGPMYDVEKLVTRTRSCRRSPKANPASGEEGWKGDShRSSPRNPTQELRRPSK SMDKKEDRGPEQESHQAQAAKAKDLVEVLPTVEEGLREVKKDTRPMSRSKHGGWLL REHQAGFEKLRRTRGEEKEAEKEKKPCMSGRRMTLRDDQPAKLEKEPKTRPEENKPE RPSGRKPRPMGIIAANVEKHETGRVIGDGNFAVVEKCRHRETRQAYAMKIIDKSLK GKEDMVDSEILIIQSLSHPNIVKLHEVYETDMEIYILILEYVQGGDLFDIAIESVKFPE PDAALMINDLCKALVHMHDKSI VHRDLKPENLLVQREDKSTTLKLADFLAKHVVRP IFTVCGTPTYVAPEILSEKGYGLEVDMWAAGVILYILLCGFPFRSPERDQDELFNII QLGHFEFLPPYWDNISDTAAKDLVSRLLVDPKKRYTAHQVLQHPWIETAGTNTVIR QKQVSPSSEGHFRSQHKRVVEQVS

	SEQ ID NO: 55	3016 bp
<p>NOV21c, CG57109-03 DNA Sequence</p>	<p>TTTACAGAGTCAGGCTCACCGTGAGAGGGCTCCTGTATTAGTCCCTTTTCATGCTGC TTATAGAGACATACCTGAGACTGGGCAATTTGCAGAGAAAGGTTTCTTGGACTTACA GTTAGTTCCACGTGGCTGGGGAAGCCTCACAATCATGGCGGAAGGCAGGAAGGGCAA GTCCCATCTTACATGGATGGCAGCAGGCAAAGAGAGAATGAGGAAGATGCAAAAGCGG AAACCCCTGATGTAACCATCAGATCTTATGAGATTTATTCACTACCATGGAACAGACA GCAAGGCGTATGTGACCATTCTCTAGAATATTTAAGCTCGAGAATCTCAGAGCGGAAG CTGCAAGGCTCCTGGCTGCCAGCCGAGGGGAATCTGGAGAAACATTCTTGGGGC CGCGTGGCCCCGTGTCCTTGTCTGCCCTCGGAATGGCTTCACTCAGCACATCC TGAGAACAGCCCTCTGAAGCCAGGGTCTGACCGTAGTGAAGCTGGGTGGGACGCG CCCCGAAAGATCACTCTGCTCCTCAACAGGCGATCAGTGCAGACGTTGAGCAGCTCT TAGCTGACATCTCAGAAGCCTTGGGCTCTCCAGATGGAAGAATGACCGTGTGAGGAA ACTGTTTAACTCAAGGGCAGGGAATCAGGAGCGTCTCTGATTCTTCAAGGAAGGG GATGCTTTCATAGCTATGGGCAAAGAACCCTGACACTGAAGAGCATTAGGTGGCTG TAGAAGAACTGTACCCCAACAAAGCCCGGGCCCTGACACTGGCCAGCAGCCGTGC CCCTTCTCCAAGGCTGAGGAGCAGGCTGTTAGCAAGGCTCTGAAAGGAGACCACCGC TGTGGGAGAGCCGAGACCCCAAGAGCTGCAGCGAAGTTGCAAGGATGCAAGGCACTA TGAGGCACAGGGGAGATCCCCGAGGAGCTTCACTAGATGACAGAGCGAGGACCCA GAAGAAGTGGGGAGGGGAAATGGGAGCCAGAACCAGTAGCAAGCCCCCAGGGAA GCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAAGCATCTTGGGGTGGAGATTGAAA AGACCTCGGGTGAAATTATCAGATGCGAGAAGTGCAAGAGAGAGAGGGAGCTTCAGCA GAGCCTGGAGCGTGAGAGGCTTCTCTGGGAGCAGTGAAGTATGGGGAAGGGC CCAATGTATGATGTGAGAGCTGGTGAAGACAGAAGCTGCAGGAGGTCTCCCGAGG CAATCCTGCAAGTGGGAGGAAGGTTGAAGGTGACAGCCACAGGAGCAGCCCGAG GAATCCCACTCAAGAGCTGAGGAGACCCAGCAAGAGCATGGACAAGAAAGAGGACAGA GGCCAGAGGATCAAGAAAGCCATGCTCAGGGAGCAGCCAGGCGCAGAGAGGACCTTG TGAAGTTCTTCTGTCTCAGAGAGGAGGCTGAGGGAGGTGAAGAAGGACACAGGCC CATGAGCAGGAGCAACATGGTGGCTGGCTCCTGAGAGAGCACCAGGCGGGCTTTGAG AAGCTCCGAGGACCCGAGGAGAGAGAAGGAGGCGAGAGAAGGAGAAAAGCCATGTA TGTCTGGAGGCAAGAAGATGACTCTCAGAGATGACCAACCTGCAAGCTAGAAAAGGA GCCAAGACGAGGCCAGAAGAGAACAGCCAGAGCGGCCAGCGGTGCGAAGCCACGG CCCATGGGCATCATTGCCGCCAATGTGGAAGCATTATGAGACTGGCCGGGTCTATTG GGGATGGGAACCTTGTCTGCTGAGGAGTGCAGACACCGGAGACAGGCAGGCGCTA TGCGATGAAGATCATTGACAAGTCCAGACTCAAGGGCAAGGAGGACATGGTGGACAGT GAGATCTTGATGAAGTCTACGAAACAGACATGGAATCTACCTGATCCTGGAGTACG TGCAAGGAGGAGACCTTTTGTAGCCATCATAGAAAGTGTGAAGTTCCCGGAGCCCGA TGCTGCCCTCATGATCATGGAATTATGCAAGCCCTCGTCCACATGCACGACAAGAGC ATTGTCCACGGGACCTCAAGCCGGAAGAACCTTTTGGTTGAGCAAAATGAGGACAAAT CTACTACCTTGAAATGGCTGATTTTGGACTGCAAGCATGTGGTGAGACCTATATT TACTGTGTGGGACCCCACTTACGTAGCTCCGAAATCTTTCTGAGAAAGGTTAT GGACTGGAGGTGGACATGTGGGCTGCTGGCGTGATCCTCTATATCTGCTGTGTGGCT TTCCCCATTCCGAGCCCTGAGAGGGACAGGACGAGCTCTTTAACATCATCCAGCT GGGCCACTTTGAGTTCCTCCCCCTTACTGGGACAAATATCTCTGATGCTGTAAAGAT CTGGTGAGCCGGTTGCTGGTGGTAGACCCCAAAAGCGCTACACAGCTCATCAGGTTT TTACGACCCCTGGATCGAAACAGCTGGCAAGACCAATACAGTGAACAGCAGAGAAGCA GGTGTCCCCAGCAGCGAGGGTCACTCCGAGCCAGCACAAGAGGGTGTGGAGCAG GTATCATAGTACCACCTTGGGAATCTGTCCAGCCCCAGTTCTGCTCAAGGACAGAG AAAAGGATAGAAGTTTGAAGAGAAAACAATGAAAGAGGCTTCTTACATAATTGGTGA ATCAGAGGGAGAGACACTGAGTATATTTAAAGCATATTAAGCAATTAAGTCAATGT TAAATGTACAACATATTTTAGATTGTATATTTAAAGCCTTTAATACATTTTGGG GGGTAAGCATTGTCTCAGTGAAGATTTTGGTAATAATGATGTGTTTTGCTTCCCT TTGTAACCAAGTTTATTCTGTACTACAGGAGTGGTCTTACAGGCTTAACTCCCC CTGTGAGATTAATAAGGTGCACTGTGGTCTTCTGTGTTAATAAATGTGCTCTGAAT</p>	
	ORF Start: ATG at 52	ORF Stop: TAG at 2617
	SEQ ID NO: 56	855 aa MW at 97447.3kD
<p>NOV21c, CG57109-03 Protein Sequence</p>	<p>MLLIETYLRNLQKRVFLDLQVPRGWGSLTMAEGKEGQVPSYMDGSRQRENEEDA KAETPDVTIRSYEIIYSLPWNRQGVCDHSLLEYLSRISERKLQGSWLPASRGNLEKPF LGPRGPVVPFLFCPRNLGSAHPENSPLKPRVVTVKLGGRPRPKITLLNRRSVQTFE QLLADISEALGSPRWKNDVRKLFNLKGREIRSVSDFREGDAFIAMGKEPLTLKSIQ VAVEELYPNKARALTLAQHSRAPSPRLRSRLFSKALKGDHRCGETETPKSCSEVAGCK AAMRHQGIPEELSLDDRARTQKKGGRGWEPSPSKPPREATLEERHARGKEHLGVE IEKTSGEIIRCEKCKRERELQQSLERERLSLGTSELDMGKGPMDVEKLVTRSCRSS PEANPASGEEGNKGDSSRRSPRNPTQELRRPSKSMDDKEDRGPEQESHAQGAAKAKK DLVEVLPTVEEGLREVKKDTRPMSRSKHGGWLLREHQAGFEKLRRTRGEEKEAEKEKK PCMSGRRMTLRDDQPAKLEKEPKTRPEENKPERPSGRKPRPMGIIAANVEKHETGR VIGDGNFAVVKECRHRETRQAYAMKIIDKSLKKGEDMVDSEILHEVYETDMEIYLIL EYVQGGDLFDALIESVKFPPEPDALMIMDLCKALVHMHDKSIIVHRDLKPNLLVQRNE DKSTTLKLADFLAKHVVRPIFTVCGTPTYVAPEILSEKGYGLEVDMAAGVILYILL CGFPPFRSPERDQDELNIILQGHFEFLPPYWDNISDAAKDLVSRLLVDPKRYTAH</p>	

	QVLQHPWIETAGKTNIVKRQKQVSPSSEGHFRSQHKRVVEQVS		
	SEQ ID NO: 57	2433 bp	
NOV21d, CG57109-04 DNA Sequence	TTTACAGAGTCAGGCCTCACCGTGAGAGGGCTCCTGTATTAGTCCCTTTTCATGCTGC TTATAGAGACATACCTGAGACTGGGCAATTTGCAGAGAAAGGTTTTCTTGGACTTAC GTTAGTTCCACGTGGCTGGGGAAGCCTCACATCATGGCGAAGGCAAGGAAGGGCAA GTCCCATCTTACATGGATGGCAGCAGGCAAGAGAGAATGAGGAAGATGCAAAAGCGG AAACCCCTGATGTAACCATCAGATCTTATGAGATTTATTCATACCATGGAACAGACA GCAAGGCGTATGTGACCATCTCTAGAATATTTAAGCTCGAGAATCTCAGAGCGGAAG CTGCAAGGCTCCTGGCTGCCAGCCGAGGGGAATCTGGAGAAACCATTCCTGGGGC CGCGTGGCCCCGTGCTGCCCTTGTCTGCCCTCGGAATGGCTTCCTCAGCACATCC TGAGAACAGCCCTCTGAAGCCCAGGGTCGTGACCGTAGTGAAGCTGGGTGGGACAGCGC CCCCGAAAGATCACTCTGCTCCTCAACAGGCGATCAGTGCAGACGTTCCAGCAGCTCT TAGCTGACATCTCAGAAGCCTTGGGCTCTCCAGATGGAAGAATGACCGTGTGAGGAA ACTGTTTAACTCAAGGGCAGGGAATCAGGAGCGTCTCTGATTTCTTCAGGGAAGGG GATGCTTTCATAGCTATGGGCAAAAGAACCTGACACTGAAGAGCATTTCAGGTGGCTG TAGAAGAACTGTACCCCAACAAAGCCCGGGCCCTGACACTGGCCAGCACAGCCGTGC CCCTTCTCAAAGGCTGAGGAGCAGGCTGTTTAGCAAGGCTCTGAAAGGAGACCACCGC TGTGGGGAGACCGAGACCCCCAAGAGCTGCAGCGAAGTTGCAGGATGCAAGGCAGCCA TGAGGCACCAAGGGGAAGATCCCCGAGGAGCTTCACTAGATGACAGAGCGAGGACCCCA GAAGAAGTGGGGGAGGGGAAATGGGAGCCAGAACCCAGTAGCAAGCCCCCAGGGAA GCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAAGCATCTTGGGGTGGAGATTGAAA AGACCTCGGGTGAAATTATCAGATGCGAGAAGTGCAAGAGAGAGAGGAGCTCCAGCA GAGCTGGAGCGTGAGAGGCTTTCTCTGGGGACCACTGAGCTGGATATGGGGAAGGGC CCAATGTATGATGTGGAGAAAAAGCCATGTATGTCTGGAGGCAGGAAGATGACTCTCA GAGATGACCAACCTGCAAGACTAGAAAAGGAGCCCAAGACGAGGCCAGAAGAGAACA GCCAGAGCGGCCAGCGGTGGAAGCCACGGCCCATGGGCATCATTTGGCCCAATGTG GAAAAGCATTATGAGACTGGCCGGGTCTTGGGGATGGGAACCTTGTGCTGTGAAGG AGTGACAGACACCGCAGAGACCAGGCAGGCTATGCGATGAAGATCATTGACAAGTCCAG ACTCAAGGGCAAGGAGGACATGGTGGACAGTGAGATCTTGATCATCCAGAGCCTCTCT CACCCCAACATCGTGAATTTGCATGAAGTTTACGAAACAGACATGGAATCTCACTGA TCCTGTGAGTACGTGCAGGAGGAGACCTTTTTCAGCCCATCATAGAAAGTGTGAAGTT CCCGGAGCCCGATGCTGCCCTCATGATCATGGACTTATGCAAGGCCCTCGTCCACATG CAGCAACAAGAGCATTTGTCCACCGGGACCTCAAGCCGGAACCTTTTGGTTTCAGCGAA ATGAGGACAAATCTACTACCTTGAATTTGGCTGATTTGGACTTGCAAGCATGTGGT GAGACCTATATTTACTGTGTGTGGGACCCCACTTACGTAGCTCCCGAAATCTTTCT GAGAAAGGTTATGACTGGAGGTGGACATGTGGGCTGTGGCGTGATCTCTATATCC TGCTGTGTGGCTTTCCCCATTCCGAGCCCTGAGAGGGACACGAGCAGAGCTTTTAA CATCATCCAGCTGGGCCACTTTGAGTTCTCCCCCTTACTGGGACAATATCTCTGAT ACAGCTGTCTAAAGATCTGGTGAAGCCGTTGCTGGTGGTAGACCCCAAAAGCGCTACA CAGCTCATCAGGTTCTTCAGCACCCCTGGATCGAAACAGCTGGCAAGACCAATACAGT GAAACGACAGAAGCAGGTGTCCCCAGCAGCGAGGGTCACTTCCGGAGCCAGCAAG AGGGTTGTGGAGCAGGTATCATAGTCACCACCTTGGGAATCTGTCTCAGCCCCAGTTCT TGCTCAAGGACAGAGAAAAGGATAGAAGTTTGAGAGAAAAACAATGAAGAGGCTTCT TCACATAATTGGTGAATCAGAGGGAGAGACACTGAGTATATTTAAAGCATATTA		
	ORF Start: ATG at 52	ORF Stop: TAG at 2284	
	SEQ ID NO: 58	744 aa	MW at 84729.4kD
NOV21d, CG57109-04 Protein Sequence	MLLIETYLRLGNLQKRVFLDLQLVPRGWSLTIMAEKQVPSYMDGSRQRENEEDA KAETPDVTIRSYEISLWNRQQGVCDHSLEYLSRISERKLQGSWLPASRGNLEKPF LGPRGPVPLFCPRNGLHSAHPENSLKPRVTVVVLGGQRPRIITLLNRRSVQTFE QLLADISEALGSPRWKNDVRKLFNLKGREIRSVSDFREGDAFIAMGKEPLTLKSIQ VAVEELYPNKARALTQAHSRAPSPRLSRSLFSKALKGDHRCGETETPKSCSEVAGCK AAMRHQGIPEELSLDDRARTQKWKGRGKWEPEPSSKPPREATLEERHARGEKHLGVE IEKTSGEIIRCEKCRERELQSLERERLSLGTSELDMGKPMYDVEKPKCMGSGRRM TLRDDQPAKLEKEPKTRPEENKPERPSGRKPRPMGI AANVEKHETGRVIGDGNFAV VKECRHRETRQAYAMKIIDKSRLKGKEDMVDSEILIIQSLSHPNIVKLHEVYETDMEI YLILEYVQGGDLFDALIESVKFPEPDALMIMDLCKALVHMHDSIVHRDLKPENLLV QRNEDKSTLKLADFLAKHVVRPIFTVCGTPTYVAPEILSEKGYGLEVDMAAGVIL YILLCGFPFRSPERDQELFNIIQLGHFEFLPPYWDNISDTAAKDLVSRLLVVDPKK RYTAHQVLQHPWIETAGKTNIVKRQKQVSPSSEGHFRSQHKRVVEQVS		

	SEQ ID NO: 59	2133 bp	
NOV21e, CG57109-05 DNA Sequence	TTGACCGTGTGAGGAACTGTTTAACTCAAGGGCAGGGAAATCAGGAGCGTCTCTG TTTCTTCAGGGAAGGGGATGCTTTCATAGCTATGGGCAAAGAACCACTGACACTGAAG AGCATTACAGTGGCTGTAGAAGAACTGTACCCCAACAAAGCCCGGGCCCTGACACTGG CCCAGCACAGCCGTGCCCTTCTCCAAGGCTGAGGAGCAGGCTGTTTAGCAAGGCTCT GAAAGGAGACCACCGCTGTGGGAGACCGAGACCCCAAGAGCTGCAGCGAAGTTGCG GGATGCAAGGCAGCCATGAGGCACAGGGGAAGATCCCCGAGGAGCTTTCACATAGATG ACAGAGCGAGGACCCAGAAGAAGTGGGGAGGGGAAATGGGAGCCAGAACCAGTAGT CAAGCCCCCAGGGAAGCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAGCATCTT GGGTGGAGATTGAAAAGACCTCGGGTGAAATTATCAGATGCTGAGAAGTCAAGAGAGA AGAGGGAGCTCCAGCAGAGCTGGAGCGTGAGAGGCTTTCTCTGGGGACCACTGAGCT GGATATGGGGAAGGGCCCAATGTATGATGTGGAGAAGCTGGTGAGGACCAGAAGCTGC AGGAGGTCTCCCGAGGCAATCCTGCAAGTGGGGAGGAAGGGTGGAAAGGTGACAGCC ACAGGAGCAGCCCCAGGAATCCCACTCAAGAGCTGAGGAGACCCAGCAAGAGCATGGA CAAGAAAGAGGACAGAGGCCAGAGGATCAAGAAAGCCATGCTCAGGGAGCGGCCAAG GCCAAGAAGGACCTTGTGGAAGTTCTTCCTGTCAAGAGGAGGGGCTGAGAGAGGTG AGAAGGACACCGGCCCATGAGCAGGAGCAACATGGTGGCTGGCTCTGAGAGAGCA CCAGCGGGCTTTGAGAAGCTCCGAGGACCCGAGGAGAAGAGAAGGAGGCAGAGAAG GAGAAAAAGCCATGTATGTCTGGAGGCAGAAGGATGACTCTCAGAGACGACCAACCTG CAAGCTAGAAAAGGAGCCCAAGACGAGGCCAGAAAGAGAAACAAGCCAGAGCGGCCCA CGGTGGAAGCCACGGCCCATGGGCATCATTGCCGCCAATGTGGAAGACATTATGAG ACTGGCCGGGTCTATGGGGATGGGAACCTTTGCTGCTGTAAGGAGTGACAGACCCGCG AGACCAGGCAGGCCATATGCGATGAAGATCATTGACAAGTCCAGACTCAAGGGCAAGGA GGACATGGTGGACAGTGAGATCTTGATCATCCAGAGCCTCTCTACCCCAACATCGTG AAATTGATGAAGTCTACGAAACAGACATGGAAATCTACCTGATCTGGAGTAGCTGC AGGGAGGAGACCTTTTACGCCATCATAGAAAGTGTGAAGTCCCGGAGCCCGATGC TGCCCTCATGATCATGGACTTATGCAAGCCCTCGTCCACATGCAGCACAAGAGCATT GTCCACCGGACCTCAAGCCGGAACCTTTTGGTTCAAGCAATGAGGACAAATCTTA CTACCTTGAATTTGGCTGATTTTGGACTTGCAAGCATGTGGTGAGACCTATATTATTA TGTGTGTGGGACCCCACTTACGTAGCTCCCGAAATCTTTCTGAGAAAGGTTATGGA CTGGAGGTGGACATGTGGGCTGCTGGCGTGATCCTCTATATCCTGCTGTGTGGCTTTC CCCCATTCCGAGCCCTGAGAGGGACAGGACGAGCTCTTAACTATCATCCAGCTGGG CCACTTTGAGTTCTCTCCCCCTTACTGGGACAATATCTCTGATGCTGCTAAAGATCTG GTGAGCCGGTTGCTGGTGGTAGACCCCAAAAGCGCTACACAGCTCATCAGGTTCTTC AGCACCCTGGATCGAAACAGCTGGCAAGACCAATACAGTGAAACGACAGAAGCAGGT GTCCCCCAGCAGCGAGGGTCACTTCCGAGCCAGCACAAGAGGTTGTGGAGCAGGTA TCATAGTCACCACCTTGGGAATCTGTCCAGCCCCAGTTCTGCTCAAGGACAGAGAA AGGATAGAAGTTTGAGAGAAAAACAATGAAAGAGGCTTCTTCACA		
	ORF Start: ATG at 90	ORF Stop: TAG at 2034	
	SEQ ID NO: 60	648 aa	MW at 73813.6kD
NOV21e, CG57109-05 Protein Sequence	MGKEPLTLKSIQVAVEELYPNKARALTQAHSRAPSPRLRSRLFSKALKGDHRCGETE TPKSCSEVAGCKAAMRHQGIPEELSLDDRARTQKKWGRGKWEPEPSSKPPREATLEE RHARGEKHLGVEIEKTSGBEIRCEKCKRERELQOSLERERLSLGTSELDMGKGPYDV EKLVRTRSCRRSPKANPASGEEGWKGDHRSSPRNPQELRRPSKSMDDKEDRGPEDQ ESHAQGAAKAKKDLVEVLFPVTEBGLREVKKDTRPMSRSKHGGWLLREHQAGFEKLRR RGEKEAEKEKPCMSGGRMTLRDDQPAKLEKEPKTRPEENKPERPSGRKPRPMGI AANVEKHYETGRVIGDGNFAVVKECRHRETRQAYAMKIDKSRLKGKEDMVDSEILII QSLSPNIVKLHEVYETDMEIYLIILEYVQGGDLFDAIIESVKFPEPDAALMIMDLCKA LVHMHDKSIVHRDLKPENLLVQRNEDKSTTLKLADFGLAKHVVRPIFTVCGTPTYVAP EILSEKGYGLEVDMWAGVILYILLCGFPFRSPERDQDELFNIIQLGHFEFLPPYWD NISDAAKDLVSRLLVVDPKKRYTAHQVLQHPWIETAGKTNTVKRQKQVSPSSEGHFRS QHKRVVEQVS		

	SEQ ID NO: 61	2720 bp	
NOV21f, CG57109-06 DNA Sequence	GGCACTGACATGGACTGAAGGAGTAGAAAAGAAGCCTTGGGCTCTCCAGATGGAAAG AATGACCGTGTGAGGAAACTGTTTAACTCAAGGGCAGGGAATCAGGAGCGTCTCTG ATTTCCTCAGGGAAGGGGATGCTTTCATAGCTATGGGCAAAGAACCACCTGACACTGAA GAGCATTTCAGGTGGCTGTAGAAGAACTGTACCCCAACAAGCCCGGGCCCTGACACTG GCCCAGCACAGCCGTGCCCTTCTCCAAGGCTGAGGAGCAGGCTGTTTAGCAAGGCTC TGAAAGGAGACCACCGCTGTGGGAGACCGAGACCCCAAGAGCTGCAGCGAAGTTGC AGGATGCAAGGCAGCTATGAGGCACCAGGGGAAGATCCCCGAGGAGCTTTCCTAGAT GACAGAGCGAGGACCAGAGAAGTGGGGGAGGGGAAATGGGAGCCAGAACCAGTA GCAAGCCCCCAGGGAAGCCACTCTGGAAGAGAGGCACGCAAGGGGAGAGAAGCATCT TGGGGTGGAGATTGAAAAGACCTCGGGTGAAATTATCAGATGCGAGAAGTCAAGAGA GAGAGGGAGCTTCAGCAGAGCCTGGAGCGTGAGAGGCTTCTCTGGGACCACTGAGC TGGATATGGGGAAGGGCCCAATGTATGATGTGGAGAAGCTGGTGAGGACCAAGAAGCTG CAGGAGGTCTCCGAGGCAAACTCTGCAAGTGGGGAGGAAGGTTGGAAGGTGACAGC CACAGGAGCAGCCCCAGGAATCCCACTCAAGAGCTGAGGAGACCCAGCAAGACATGG ACAAGAAAGAGGACAGAGGCCAGAGGATCAAGAAAGCCATGCTCAGGAGCAGCCAA GGCCAAAGAAGGACCTTGTGGAAGTTCTTCTGTACAGAGGAGGGGCTGAGGGAGGTG AAGAAGGACACAGGCCCATGAGCAGGAGCAAAACATGGTGGCTGGCTCCTGAGAGAGC ACCAGGCGGGCTTTGAGAAGCTCCGAGGACCCGAGGAGAAGAGAAGGAGGCAGAGAA GGAGAAAAGCCATGTATGTCTGGAGGCAGAAAGGATGACTCTCAGAGATGACCAACCT GCAAGCTAGAAAAGGAGCCCAAGACGAGGCCAGAAGAGAACAAGCCAGAGCGGCCCA GCGGTGGAAGCCACGGCCCATGGGCATCATTCGCCCAATGTGGAAAGCATTATGA GACTGGCCGGGTCTATGGGGATGGGAACCTTGTCTGTCTGAAGGAGTGACAGACCCGC GAGACCAGGAGGCTATGCGATGAAGATCATTGACAAGTCCAGACTCAAGGGCAGGG AGGACATGGTGGACAGTGAGATCTTGATGAAGTCTACGAAACAGACATGGAAATCTA CCTGATCCTGGAGTACGTGCAGGGAGGAGACCTTTTTCAGCCATCATAGAAAGTGTG AAGTTCCCGAGCCCGATGCTGCCCTCATGATCATGACTTATGCAAAGCCCTCGTCC ACATGCACGACAAGAGCATTGTCCACCGGGACCTCAAGCCGGAACCTTTTGGTTCA GCGAATGAGGACAATCTACTACCTTGAAATGGCTGATTGTTGACTTGAAGAAGCAT GTGGTGAGACCTATATTTACTGTGTGTGGGACCCCAACTTACGTAGCTCCCGAAATTC TTTCTGAGAAAGGTAAGTGTTACACATCGATCTGTGGGACTCTAGTTCCCTTACTAAC AAATGTTTCATTGTGCATATTTACTAGTTTCAATATGGAATAAATCTCAGAGAAGT ACACTTAGGCTTGGATTGGACTTCAATGAAAATATTTGAAGTAGGCTTGAACCAAGC ATGAGAGCTTCTCTCATTAGGGCTGCCCTTGTGTACAGTCAATGGATCAGTGTGTGT GCATGTGTGTGTGTGTGTGTGTATTGTGTTTAGGCAGGACAGTGAGATGAAGA TGATGGAGAATTGGGTGGGGAACCTCAAGCAAGCAAGGCTACTTGACCAAGGCTATCT CTAATAGGAGAGAATTGAAGCAGTCCTTATGGTACTTGGTTTAAAAATTTCTTACCA ACCTTGCAATTAAAGGAAAAGGATCCCATTTTCTCCATGAAGCTCTATGAATATTTAT TACCTACTTGTATATTATGCAAGGATCCAATGAGCTGTTTTAGTGACAACTTTCTAA AACATTTAAAGGAAATAATAGTTATGATATGGCTCTAAATATATGATGACTATTT GACTACTGTGGCACTCCAGGAGAAACCAATTTACCCACCATTGGTAAGATGGGAAGAC TCTCATTGGGTTGCAGGGTTGGTGACAGGGAGAAGGGATGAGCGGATGTTTCCAGC AGCAGAAGCATCAGGATAATTAAGATGAGGAGATGGCCAGGGATGGTGGCTCATGCTC GTAATCCCAGCTCTTTGGGAGGCTGAGGCAGGTGGATCACCTGAGGTGAGGCGTTTGA AACCATCCTGGCCAACATGGTGAAACCTGTCTGTACTAAAACACTCAAAAAAATTAGC TGGGCGTGGTGGCACATGCCTGTAATCTGAGCTACTCGGGAGGCTGAGGCAGGAGAAT TGCTTGAACCGGGAGGTGGAGTTTGCAAGTGAAGCAAGATCGTGCCATTGCACTCCAG CCTGGGCAACAAGAGTGAACCTTCATCTCAAAAAAAAAAAAAAAAAAAAAAC		
	ORF Start: ATG at 149	ORF Stop: TGA at 1826	
	SEQ ID NO: 62	559 aa	MW at 63436.9kD
NOV21f, CG57109-06 Protein Sequence	MGKEPLTLKSIQVAVEELYPNKARALTLAQHSRAPSRLRSLFSLKALGKDHRCGETE TPKSCSEVAGCKAAMRHQGGKIPEELSLDDRARTQKKWGRGKWEPEPSSKPPREATLEE RHARGEKHLGVEIEKTSGEIIRCEKCKRERELQSSLERERLSLGTSELDMGKGPMDV EKLVRTRSCRSPPEANPASGEEGWKGDSSRSPRNPTQELRRPSKSMDDKEDRGPEQ ESHAQGAAKAKKDLVEVLPVTEEGLREVKKDTRPMSRSKHGGWLLREHQAGFEKLRR RGEKEAEKEKKPCMSGGRMTLRDDQPAKLEKEPKTRPEENKPERPSGRKPRPMGII AANVEKHVETGRVIGDGNFAVVKECRHRETRQAYAMKIDKSLRKGKEDMVDSEILHE VYETDMEIYLLILEVYQGGDLFDALIESVKFPEPDALMIMDLCKALVHMHDKSIVHRD LKPENLLVQRNEDKSTTLKLDLGLAKHVVRPIFTVCGTPTYVAPEILSEKGYKTSI CGTLVPLLTNVSVFVFTSFQYGINLRELTLRLGFGLO		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 21B.

Table 21B. Comparison of NOV21a against NOV21b through NOV21f.		
Protein Sequence	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV21b	1..648 188..836	608/649 (93%) 608/649 (93%)
NOV21c	1..648 221..855	595/648 (91%) 595/648 (91%)
NOV21d	300..648 395..744	349/350 (99%) 349/350 (99%)
NOV21e	1..648 1..648	608/648 (93%) 608/648 (93%)
NOV21f	1..529 1..516	476/529 (89%) 476/529 (89%)

Further analysis of the NOV21a protein yielded the following properties shown in Table 21C.

Table 21C. Protein Sequence Properties NOV21a	
PSort analysis:	0.7000 probability located in plasma membrane; 0.3000 probability located in microbody (peroxisome); 0.3000 probability located in nucleus; 0.2000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV21a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 21D.

Table 21D. Geneseq Results for NOV21a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY42696	Rat serine-threonine protein kinase PK80 sequence - Rattus norvegicus, 733 aa. [WO9950395-A1, 07-OCT-1999]	1..648 98..733	495/649 (76%) 535/649 (82%)	0.0
AAB65622	Novel protein kinase, SEQ ID NO: 148 - Mus musculus, 806 aa. [WO200073469-A2, 07-DEC-2000]	1..647 172..805	472/648 (72%) 521/648 (79%)	0.0
AAY42697	Mouse serine-threonine protein kinase PK80 sequence - Mus musculus, 732 aa. [WO9950395-A1, 07-OCT-1999]	1..647 98..731	472/648 (72%) 521/648 (79%)	0.0
AAB65621	Novel protein kinase, SEQ ID NO: 147 - Homo sapiens, 686 aa. [WO200073469-A2, 07-DEC-2000]	276..648 313..686	370/374 (98%) 371/374 (98%)	0.0
AAY54579	A rat calcium/calmodulin dependent protein kinase designated CaMK-VI - Rattus sp, 422 aa. [EP978562-A1, 09-FEB-2000]	348..641 75..367	154/294 (52%) 212/294 (71%)	3e-87

In a BLAST search of public sequence databases, the NOV21a protein was found to have homology to the proteins shown in the BLASTP data in Table 21E.

Table 21E. Public BLASTP Results for NOV21a				
Protein Accession Number	Protein/Organism/Length	NOV21a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9C098	KIAA1765 PROTEIN - Homo sapiens (Human), 608 aa (fragment).	41..648 1..608	608/608 (100%) 608/608 (100%)	0.0
O15075	Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-like and CAM kinase-like 1) - Homo sapiens (Human), 740 aa.	348..641 382..674	155/294 (52%) 212/294 (71%)	2e-87
Q9JLM6	CPG16 - Mus musculus (Mouse), 433 aa.	348..641 75..367	154/294 (52%) 212/294 (71%)	7e-87
O08875	Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-like and CAM kinase-like 1) (Calcium/calmodulin-dependent protein kinase type I-like CPG16) - Rattus norvegicus (Rat), 433 aa.	348..641 75..367	154/294 (52%) 212/294 (71%)	7e-87
Q9JLM8	Serine/threonine-protein kinase DCAMKL1 (EC 2.7.1.-) (Doublecortin-like and CAM kinase-like 1) - Mus musculus (Mouse), 756 aa.	348..641 398..690	154/294 (52%) 212/294 (71%)	7e-87

PFam analysis predicts that the NOV21a protein contains the domains shown in the Table 21F.

Table 21F. Domain Analysis of NOV21a			
Pfam Domain	NOV21a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Pkinase	356..613	107/297 (36%) 219/297 (74%)	2.3e-101

Example 22.

- 5 The NOV22 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 22A.

Table 22A. NOV22 Sequence Analysis

	SEQ ID NO: 63	3553 bp
NOV22a, CG57366-01 DNA Sequence	TACATTCTTCATCGTTTAGATCAAGAAGAAGCTTTGCGACAACACCTCACAAAAGAAA CTGCGAGAGATGTTAAATCAACTGCACATTAAAAGCAGTGGATGCTTCTTTACCTAGA ACGAGTTTTAGATGGAGTTGTAGAAAAATTTATTATGTTAAGAGAAATTCGTGACATC CCAGGAACCTCAAATGGTTTATATCTCTGGCTGTGCCAAAGACTTTTGTGAAGAAAC AATTTGCAAAGGTTTCAGCCTATTTGAATGTGATTCTTGCAGCCTGCCGACCTTTGAC CATAACGGAATTATATCAGCAGTATGGACCAGAAACATGTCGTTAACTTTGGGAAGAT TTTCAACGCAAGTTAGATATCCTCTCCAACTTCTTGTGATGGACTAGGAAATACAA AAATACTGTTTCATTATAGTTTTGCCGAGTGGCTTCTGGATGTGAACACCTGTACTCA GAAGTATTTATGTAATGCAGCAGAAGGACACAGAATGTTGGCTATGAGTTATACCTGT CAAGCCAAGAATTAACACCATTTGGAAGCACAAGAATTTGCATTGCACCTTAATTAAC CAAACTTACAATTAGAGCAGCGGAGTTAGCTCTGTGGATGATATGGAATGGTACACC TGTGAGAGATTCCCTTTCTACTTTGATACCCAAGGAACAAGAAGTGCTACAGCTGTTG GTTAAAGCTGGGGCTCATGTCAACAGTGAAGACGATCGCAGATCATGCTATGTTCCGAC AAGCCTTAGAAAGAGAGGATTCCATTCCGACATTATTAGATAATGGAGCTTCAGTAA TCAGTGTGATTCAAATGGGAGAACATTATTGGCTAATGCTGCATATAGTGGCAGCTCT GATGTAGTCAATTTACTTGTCTCTAGGGGAGCAGATTTAGAGATAGAAGATGCTCATG GACATACACCACTCACTCTAGCGGCTAGACAGGGACATACCAAGGTGGTTAATTTGTTT GATTGGGTGTGGAGCAAAATTAATCATACTGATCAAGATGGTTGGACAGCATTAAAGA TCTGCTGCTTGGGGTGGCCATACTGAGGTAGTTTCTGCACTACTTTATGCTGGCGTAA AAGTGGATTGTGAGATGCTGATAGCCGAACAGCTTTGAGAGCAGCAGCATGGGGAGG ACACGAGGATATTGTACTGAATTTGCTACAAATGCGCGCTGAAGTGAACAAAGCTGAT AATGAAGGTAGAACTGCTTTGATAGCAGCAGCATACATGGGACATAGAGAGATTGTGG AACACCTACTGGACCATGGAGCAGAAGTAAATCATGAGGATGTTGATGGCAGGATGC ACTCTCTGATGCTGCACTTTGTGTGCTGCAAGTAAAGGGCAGCATCAGTTGTTAGC CTTTAATTGATCGAGGTGCTGAAGTAGATCATTGTGATAAAGATGGCATGACTCCAC TGCTGGTAGCTGCCTATGAAGGACATGTTGATGTGGTTGACTTGCTTCTAGAGGGGG AGCAGATGTAGATCACAGATAACAATGGCCGTACACCCCTCTTAGCAGCAGCGTCT ATGGGTGATGATCAGTTGTAATAACACTTTTGTGTTGGGGTGCAGCTGTGGATAGTA TTGATAGTGAAGGTAGGACAGTCTCAGTATAGCTTCAGCACAAGGAAATGTTGAGT GGTACGTACTCTACTGGATAGAGGGTTAGATGAAAATCACAGAGATGATGCTGGATGG ACACCTTTGCACATGGCAGCTTTTGAAGGGCACAGATTGATATGTGAAGCACTTATTG AACAAAGGTGCTAGAACAAATGAGATTGACAATGATGGACGAATCCCTTTTCATATTAGC TTCACAAGAGGGTCATTATGATTGTGTTCAAATATTACTGGAAAACAATCCAACATT GATCAAAGAGGTTATGATGGAAGAAATGCAGTGCAGGTTGCTGCATTAGAAGGGCACA GGGACATTGTTGAATGCTTTTAGCCATGGTGTGATGTTAACTGCAAGAGATGCTGA TGGTCCGCCCTACACTTATATCTTGGCCTTAGAAAATCAGCTTACAATGGCCGAATAT TTTTTAGAAAATGGTGCAACGTAGAAGCAAGTGATGCTGAAGGAAGGACAGCACTTC ATGTGTCTTGTGGCAAGGCCATATGGAATGGTGCAGTCTGATAGCATACCATGC TGACGTCAATGCTGCAGACATGAAAAGCGCTCTGCTTTGCAGTCTGCAGCCTGGCAG GGCCATGTAAAAGTGGTTGAGCTTCTGATTGAGCATGGTGTGATGTTGAGCATCAT GTAACCAAGGTGCACTGCACTCTGATTGTCAGCCAGGAAGGGCACATTGATGTTGT TCAGGTCTTATTAGAGCATGGTGTGATCCAAACCATGCTGATCAATTTGGACGCACT GCTATGCGGTGTTGCAGCCAAAAATGGACATTCTCAGATAAATAATTATTAGAAAAAT ATGGTGCATCTAGTTTGAATGGCTGTTCCCATCTCCTGTTTACACAAATGGAGCAAAA ACCTCTACAGTCATTGTCTTCAAAGTGCAGTCATTAACAATTAAATCAAATAGCTCT GGTAGTACTGGTGGAGGGGATATGCAGCCTTCGTTACGTGGTTTACCTAATGGGCCTA CTCATGCTTTTAGTTCTCCTTCAGAATCTCCAGATTCTACAGTTGACCGGCAGAGTC ATCACTGTCAAATAATTCCCTGAAAAGCTCAAAAAATTCATCTTTGAGAACTACTTCA TCTACAGCAACGGCTCAAACAGTGCCAATTGATAGCTTTCATAACTTGTCAATTTACAG AACAAATTCAGCAGCATTTCATTGCCACGCAGTAGAAGTCGACAGTCAATTTGTTCCCC ATCTTCCACAACACAGTCTTAGGACAGAGTCATAATTCACCAAGTAGTGAATTTGAG TGGAGTCAAGTAAAGCCAGTTTGAAGTCAACTAAAGCAAGTAAAGGGGGGAAATCAG AAAAATCTGCAAGTCTGGATCAGCTGGGAAAAAAGCGAAACAAAGTAATTTCTTACA GCCAAAGGTTTAGAATATGAAATGACTCAGTTTGTAGAGAGGACCTATAGCCAAA TCCGGGACTGCTGCACCACTAAACAAATGCCAGCAGAATCTCAATGCAAAATTTATGA TACCTTCAGCTCAGCAGGAAATGGTGCATCTCAACAGCAGTTTCTTATTCACCAACA AAGTGGGGAACAGAAGAAGAGAAATGGAATAATGACAAATCCAAATTCATCTTCAG AGCAACCAAGTTTCTTGGTAGGGTTTTCAGTCCCACGAACAAATGCAAGTAGAGGGC ATCAGGAAGTGTGGAGGGATACCCCTTCTCAGAGACAGAATTAAGCCTTAAACAGC TCTGAAGCTTCAGATTGAAGGTTCTGACCTAGCTTCAACTATAAAAGGAAACACCA TTATAAAGGTAATATTTTGTCAACATAAAGAGTAAAAACAAAGTAGAGTTCCTCTC TACTTGGATGCTCT	

	ORF Start: ATG at 67	ORF Stop: TAA at 3484	
	SEQ ID NO: 64	1139 aa	MW at 123942.1kD
NOV22a, CG57366-01 Protein Sequence	MLNQLHIKSSGCFLYLRLVDGVVENFIMLRREIRDIPGTLNGLYLWLCQRLFVRKQFA KVQPIILNVILAACRPLTITELYHAVWTKNMSLTLEDFQRKLDILSKLLVDGLGNTKIL FHYSFAEWLLDVKHCTQKYLCAAEGRMLAMSYTCQAKNLTPLEAQEFALHLINSNL QLETAELALWMIWNGTFVRDSLSTLIPKEQEVQLLVKAGAHVNSEDDRTSCIVROAL EREDSIRTLLDNGASVNCDSNGRLLANAAYSGLDVVNLLVSRGADLEIEDAHGHT PLTLAARQGHTKVVNCLIGCGANINHTDQDGWTALRSAAWGGHTEVVSALLYAGVKVD CADADSRTALRAAAWGGHEDIVLNLQHGAEVKNADNEGRALIAAAYMGHREIVEHL LDHGAEVNHEVDGRTALSVAAALCVPASKGHASVVSLIDRGAEDVHCDKDGMTPLLV AAYEGHVDVVLLLEGGADVVDHTDNNGRTPLLAAASMGHASVVNTLLFWGAAVDSIDS EGRTVLSIASAQGNVEVVRTLLDRGLDENHRDDAGWTPMHMAAFEGHRLICEALIEQG ARTNEIDNDGRIPFILASQEGHYDCVQILLNKSNIQRGYDGRNALRVALEGHRI VELLFSGADVNCCKDADGRPTLYILALENQLTMAEYFLENGANVEASDAEGRTALHVS CWQGHMEMVQVLIAYHADVNAADNEKRSALQSAWQGHVQVQLLIEHGAVVDHTCNQ GATALCIAAQEGHIDVVQVLLLEHGADPNHADQFGRGTAMRVAAKNGHSQIIKLEKYGA SSLNGCSPSPVHTMEQKPLQSLSSKVQSLTIKSNSSSGSTGGDMQPSLRGLPNGPHTA FSSPSPSPDSTVDKQSSLSNLSKSSKNSLRRTTSSTATAQTVPIDSFHNLSPTEQI QQHSLPRSRSRQSIIVSPSSTTQSLGQSHNSPSEFEWSQVKPSLKSTKASKGKSENS AKSGSAGKKAKQSNSSQPKVLEYEMTQFDRRGPIAKSGTAAPPKQMPAESQCKMIPS AQQEIGRSQQQLFIHQSGEQKRNIGMTNPNYHLQSNQVFLGRVSVPRMQDRGHQE VLEGYPSSETELSLKQALKLQIEGSDPSFNYKKTPL		

Further analysis of the NOV22a protein yielded the following properties shown in Table 22B.

Table 22B. Protein Sequence Properties NOV22a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV22a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded
 several homologous proteins shown in Table 22C.

Table 22C. Geneseq Results for NOV22a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV22a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU20535	Human secreted protein, Seq ID No 527 - Homo sapiens, 447 aa. [WO200155326-A2, 02-AUG-2001]	699..1139 7..447	439/441 (99%) 440/441 (99%)	0.0
ABB64823	Drosophila melanogaster polypeptide SEQ ID NO 21261 - Drosophila melanogaster, 2119 aa. [WO200171042-A2, 27-SEP-2001]	2..1079 885..2074	413/1213 (34%) 621/1213 (51%)	e-177
AAU25462	Human mddt protein from clone LI:334386.1:2000MAY01 - Homo sapiens, 256 aa. [WO200162922-A2, 30-AUG-2001]	466..702 2..238	231/237 (97%) 233/237 (97%)	e-132
AAU20651	Human secreted protein, Seq ID No 643 - Homo sapiens, 182 aa. [WO200155326-A2, 02-AUG-2001]	661..818 1..158	156/158 (98%) 156/158 (98%)	7e-86
ABB67412	Drosophila melanogaster polypeptide SEQ ID NO 29028 - Drosophila melanogaster, 1549 aa. [WO200171042-A2, 27-SEP-2001]	236..813 215..785	196/578 (33%) 308/578 (52%)	2e-83

In a BLAST search of public sequence databases, the NOV22a protein was found to have homology to the proteins shown in the BLASTP data in Table 22D.

Table 22D. Public BLASTP Results for NOV22a				
Protein Accession Number	Protein/Organism/Length	NOV22a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9ULJ7	Hypothetical protein KIAA1223 - Homo sapiens (Human), 768 aa (fragment).	372..1139 1..768	768/768 (100%) 768/768 (100%)	0.0
AAH24725	KIAA1223 PROTEIN - Homo sapiens (Human), 743 aa.	397..1139 1..743	743/743 (100%) 743/743 (100%)	0.0
AAL89945	SD03956P - Drosophila melanogaster (Fruit fly), 1282 aa.	2..1079 48..1237	413/1213 (34%) 621/1213 (51%)	e-177
AAL39916	SD01389P - Drosophila melanogaster (Fruit fly), 2119 aa.	2..1079 885..2074	413/1213 (34%) 621/1213 (51%)	e-177
Q9VAU5	CG10011 PROTEIN - Drosophila melanogaster (Fruit fly), 2119 aa.	2..1079 885..2074	413/1213 (34%) 621/1213 (51%)	e-177

PFam analysis predicts that the NOV22a protein contains the domains shown in the Table 22E.

Table 22E. Domain Analysis of NOV22a			
Pfam Domain	NOV22a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Ank	254..286	16/33 (48%) 26/33 (79%)	2.4e-06
Ank	287..319	17/33 (52%) 27/33 (82%)	9.1e-09
Ank	320..352	17/33 (52%) 23/33 (70%)	7.1e-06
Ank	353..385	16/33 (48%) 23/33 (70%)	2.6e-06
Ank	386..418	17/33 (52%) 26/33 (79%)	4.5e-09
Ank	419..456	17/38 (45%) 31/38 (82%)	2.9e-05
Ank	457..489	20/33 (61%) 30/33 (91%)	1.4e-10
Ank	490..522	17/33 (52%) 22/33 (67%)	6.2e-06
Ank	523..555	16/33 (48%) 27/33 (82%)	3.9e-06
Ank	556..588	15/33 (45%) 24/33 (73%)	9e-08
Ank	589..621	13/33 (39%) 26/33 (79%)	7.3e-06
Ank	622..654	17/33 (52%) 26/33 (79%)	3e-08
Ank	655..687	13/33 (39%) 23/33 (70%)	0.036
Ank	688..720	16/33 (48%) 27/33 (82%)	8.4e-10
Ank	721..753	14/33 (42%) 26/33 (79%)	2e-05
Ank	754..786	17/33 (52%) 29/33 (88%)	2.1e-10

Example 23.

The NOV23 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 23A.

Table 23A. NOV23 Sequence Analysis

	SEQ ID NO: 65	3321 bp
NOV23a, CG57368-01 DNA Sequence	TAACCCGCGCGCCCGCCCTAGCCAGCCCCGGGGCTCGGGGCTGGGGAGATCATGGCCC GCCTCTTCAGCCCCCGGCCGCCCCCAGCGAAGACCTCTTCTACGAGACCTACTACAG CCTGAGCCAGCAGTACCCGCTGCTGCTGCTGCTGCTGGGATCGTGCTCTGTGCGCTC GCGGCGCTGCTCGCAGTGGCCTGGGCCAGCGGCAGGAGCTGACCTCAGACCCGAGCT TCCTGACCACTGTGCTGTGCGCGCTGGGCGGCTTCTCGCTGCTGCTGGGCTCGCTTC CCGGGAGCAGCGACTGCAGCGCTGGACGCGTCCCCTGTCCGGCTTGGTATGGGTGCGG CTGCTAGCGCTAGGCCACGCTTCTCTGTTACCGGGGGCGTGGTGAGCGCTGGGACC AGGTGTCTTATTTTCTCTTCTGTCATCTTACGCGGTATGCCATGCTGCCCTTGGGCAT GCGGGACGCGCCGCTGCGGGGCTCGCCTCTCACTCTCGCATCTGCTGTGCTCGGG CTGTATCTTGGGCCACAGCCGACTACGCGCTGCACTGCTGCGCGAGTTGGCAGCAA ACGCAGTGCTGTTCTGTGCGGGAACGTGGCAGGAGTGTACCACAAGGCGCTGTATGGA GCGCGCCTGCGGGCCACGTTCCGGGAGGCACTCAGCTCCCTGCACTACGCGCGGCGG CTGGACACCGAGAAGAAGCACCAGGAACACCTTCTCTGTCCATCCTTCTGCTACCT TGGCCCGAGAGATGAAGGAGAGATCATGGCAGCGCTGCAGGCAGGACAGGGGTACAG GCCAGAGAGCACTAACAATTCCACAGCCTCTATGTCAAGAGGCACAGGAGTCAAG GTGCTGTATGCTGACATCGTGGGCTTACGCGGCTGGCCAGCGAGTGTTCCTTAAGG AGCTGGTGCTCATGCTCAATGAGCTCTTGGCAAGTTCGACAGATTGCCAAGGAGCA TGAATGCATGCGGATCAAGATCCTGGGGGACTGTTACTACTGTGTCTCTGGGCTGCCA CTCTCACTGCCAGACCATGCCATCACTGCGTGCGCATGGGCTGGACATGTGCGCGG CCATCAGGAACTGCGGGCAGCCACTGGCGTGGACATCAACATGCGTGTGGGCGTGCA CTCAGGCAGCGTACTGTGTGGAGTCACTGGGCTGCAGAAAGTGGCAGTACGACGTTTGG TCACATGATGTCACTGGCTAACCACATGGAGGCAGGCGGTGTACCAGGCGGAGTGC ACATCAGAGGGCTACCTGGCCCTGCTGGCAGGGGCTTATGCTGTGGAGGACGCAAG CATGGAGCATCGGGACCCCTACCTTCGGGAGCTAGGGGAGCCTACCTATCTGGTCACT GATCCACGGGCAGAGGAGGAGGATGAGAAGGGCACTGCAGGAGGCTTGTCTCTCTCGC TTGAGGGCCTCAAGATGCGTCCATCACTGCTGATGACCCGTTACCTGGAGTCTTGGG CGCAGCCAAGCCTTTTGGCCACCTGAGCCACGGAGACAGCCCTGTGTCCACCTCCACC CCTCTCCCCGCTCTCACTGTGAGGAGCCGTACCCCCGGGGACTAGATGATGAAC TGGACACCGGGGATGCCAAGTTCTTCCAGGTCACTGAGCAGCTCAACTCGCAGAAACA GTGGAAGCAGTCGAAGGACTTCAACCACTGACACTGTACTTCAGAGAGAAGGAGATG GAGAAAGAGTACCGACTCTCTGCAATCCCCGCTTCAAATACTATGAAGCCTGCACCT TCCTGGTTTTTCTCTCAACTTCACTATCCAGATGCTAGTGACAAACAGGCCCCCAGC TCTGGCCATCAGTATAGCATCACCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT TTCTCAGAGGACCTGATGAGGTGTCTCTGAAAGGCCCAAGATGCTGCACTGGCTGC CTGCACTGTCTGGCTGGTGGCCACACGACAGGACTGAGAATAGCCTTGGGCACCGC CACCATCCTCTCTTGTCTTGGCATGGCCATTACAGCCTGTTCTTCTTCCCAACATCA TCAGACTGCCCTTTCCAAGCTCCCAATGTGTCTCTCATGATTTTCAACCTCTCTCTGGG AGCTCCCTGGGTCTCTGCTCTCATCAGTGTCCCATATCCATGCACTGCTGACCGCT GGGCTTCT CTCTGCTGTGGCTGGCGGCATCCTGCTCCCTCTTCTGCACTGAGCTTCCAGCTGAAGCTGTGCTG CGGAATGCCCTCATCGTCCGCTCTATCTGGGCCCCCTGGACTCCAGGCCCGGAGTGTCT GAAGGAGCCCAAATGATGGGTGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT GTCTGGCTCGCCAGAATGAGTACTACTGCGCCTGGACTTCTGTGGAAGAAGAAGC TGAGGCAGGAGAGGAGGAGACAGAGACGATGGAGAACCTGACTCGGCTGCTCTTGGGA GAACGTGCTCCCTGCACACGTGGCCCCCAGTTTCTTGGCCAGAACCGCGCAACGAG GATCTCTACCAAGTCTATGAATGCGTTTGTGTCTCTCTGCTCAGTCCAGACT TCAAGGAGTTCTACTCTGAATCCAACATCAATCATGAGGGCCTAGAGTGTCTGAGGCT GCTCAATGAGATAATTGCTGATTTTGTATGAGCTGCTCTCAAGCCCAAGTTCACTGGG GTGGAGAAGATCAAGACCATCGGCAGCACCTACATGGCAGCCACAGGCTTAAATGCCA CCTCTGGACAGGATGCACAACAGGATGCTGAACGAGCTGCAGCCACCTTGGCACTAT GGTGGAATTGCGGTGGCCCTGGGGTCTAAGCTGGACGTCAACAAGCATTCTTCT AACAACTTCCGCTGCGAGTGGGGTGAACCATGGACCCGTAAGTCTGAGTTATTG GGGCCCAAGCCCAATATGACATTGCGGCAACACAGTGAACGTGGCCAGCCGCTAT GGAGAGTACAGGAGTCTTGGCAAAATCAAGTGAAGTGAAGAGACAGCATGGGCCCCA CAGTCCCTGGGCTACACCTGCTACAGCCGGGTGTCTCAAGGTGAAGGCAAGGGC AGCTCTGCACCTACTTCTGAACACAGACTTGACACGAAGTGGACCTCCTTCACTAC CCTAGGCTGAGATTGCACTCGCTTCTAAGAACCTCAATAAAGAGACTCTGGGGTGTCT TGGAGCCCAATTGATG	

	ORF Start: ATG at 52	ORF Stop: TGA at 3256	
	SEQ ID NO: 66	1068 aa	MW at 118800.7kD
NOV23a, CG57368-01 Protein Sequence	MARLFSRPPPPSEDLFYETYYSLSQQYPLLLLLLGLVLCALAALLAVAWASGRELTSD PSFLTTLVLCALGGFSLLLGLASREQLRQWRTPLSGLVWVALLALGHAFLTGGVVSA WDQVSYFLFVIIFTAYAMPLPLGMRDAAVAGLASSLSHLLVLGLYLGPQDPSRPALLPQL AANAVLFLCGNVAGVYHKALMERALRATFREALSLSHSRRRLDTEKKHQEHLILLSILE AYLAREMKAEMARLQAGQGSRPESTNNFHSLYVKRHQGVSVLYADIVGFTRLASECS PKELVLMNLNLFQKFDQIAKEHECMRIKILGDCYCYVSGPLSLPDHAINCVRMGLDM CRAIRKLRAATGVDINMRVGVHSGSVLCGVIGLQKWQYDVVSHDVTLANHMEAGGVPG RVHITGATLALLAGAYAVEDAGMEHRDPYLRELGEPTYLVLDIPRAEEDEKGTAGGLL SSLEGLKMRPSLLMTRYLESWGAAPFAHLSHGDSPVSTSTPLPWLSLCSRTPRGLD DELDTGDAKFFQVIEQLNSQKQWKQSKDFNPLTLTYFREKEMEKEYRLSAIPAFKYIEA CTFLVFLSNFIQMLVTNRPPALAITYSITFLLFLLILFVCFSEDLMRCLVKGPKMLH WLPALSGLVATRPLGLRIALGTATILLVFAMAITSLFFFTSSDCPFQAPNVSSMISNL SWELPGSLPLISVPYSMHCCTLGFLSCSLFLHMSFELKLLLLLWLAASCSFLHSHA WLSECLIVRLYLGLDSRPGVLKEPKLMGAISFFIFFFTLLVLARQNEYCYRLDFLWK KKLRQEREETETMENLTRLLENVLPAPVAPQFIGONRRNEDLYHQSYECVCLFASV PDFKEFYSESNINHEGLECLRLNEIIADFDELLSKPKFSGVEKKTIGSTYMAATGL NATSGQDAQQDAERSCSHLGTMTVEFAVALGSKLDVINKHSFNNFRLRVGLNHGPPVAG VIGAQPQYDIWNTVNVASRMESTGVLGKIQTVEETAWALQSLGYTCYSRGVIKVKG KGQLCTYFLNTDLTRTGPPSATLG		
	SEQ ID NO: 67	995 bp	
NOV23b, CG57368-02 DNA Sequence	<u>ATCATGGCCCGCCTCTTCAGCCCCCGCGCGCCCCCAGCGAAGACCTCTTCTACGAGA</u> <u>CCTACTACAGCCTGAGCCAGCAGTACCCGCTGCTGCTGCTGCTGCTGGGGATCGTGCT</u> <u>CTGTGCGCTCGCGGCGTCTGCGAGTGGCCTGGGCCAGCGGCAGGGAGCTGACCTCA</u> <u>GACCCGAGCTTCTTGACCACTGTGCTGTGCGCGCTGGGCGGCTTCTCGTGTGCTGG</u> <u>GCCTCGCTTCCCGGGAGCAGCGACTGCGACGCTGGACGCGTCCCTGTCCGGCTTGGT</u> <u>ATGGGTGCGCGCTAGCGCTAGGCCACGCCTTCTGTTCACCGGGGGCGTGGTGAGC</u> <u>GCCTGGGACCAAGTGTCTATTTTCTCTTCGTCATCTTCACGGCGTATGCCATGCTGC</u> <u>CCTTGGGCATGCGGGACGCCGCGCTCGCGGGCCTCGCCTCCTCACTCTCGCATCTGCT</u> <u>GGTCTCTCGGGCTGTATCTTGGGCCACAGCCGGA</u> <u>CTCAGCGCTGCACTGCTGCCGCGAG</u> <u>TGGCAGCAAACGAGTGTCTTCTGTGCGGGAACGTGCGCAGGAGTGACCAAGG</u> <u>CGCTGATGGAGCGCGCCTCGCGGGCCACGTTCCGGGAGGCACTCAGCTCCCTGCACTC</u> <u>ACGCCGGCGGCTGGACACCGAGAAGAAGCGCCAGGAACACCTTCTCTGTCCATCCTT</u> <u>CCTGCCTACTGGCCCCGAGAGATGAAGGCAGAGATCATGGCACGGCTGCAGGCAGGAC</u> <u>AGGGGTACAGGCCAGAGAGCACTAACAATTTCCACAGCCTCTATGTCAAGATACAGG</u> <u>AGTCCTTGGCAAATCCAAGTGA</u> <u>CTGAGGAGACAGCATGGGCCCTACAGTCCCTGGGC</u> <u>TACACCTGTACAGCCGGGGTGTCA</u> <u>TCAAGGTGAAAGGCCAAGGGCAGCTCTGCACCT</u> <u>ACTTCTGAACACAGACTTGACACGAACCTGACCTCCTTCAGCTACCTCAGGCTGAGA</u> <u>TTGCACTCG</u>		
	ORF Start: ATG at 4	ORF Stop: TGA at 982	
	SEQ ID NO: 68	326 aa	MW at 35652.1kD
NOV23b, CG57368-02 Protein Sequence	MARLFSRPPPPSEDLFYETYYSLSQQYPLLLLLLGLVLCALAALLAVAWASGRELTSD PSFLTTLVLCALGGFSLLLGLASREQLRQWRTPLSGLVWVALLALGHAFLTGGVVSA WDQVSYFLFVIIFTAYAMPLPLGMRDAAVAGLASSLSHLLVLGLYLGPQDPSRPALLPQL AANAVLFLCGNVAGVYHKALMERALRATFREALSLSHSRRRLDTEKKHQEHLILLSILE AYLAREMKAEMARLQAGQGSRPESTNNFHSLYVKSTGVLGKIQTVEETAWALQSLGY TCYSRGVIKVKGKQLCTYFLNTDLTRTGPPSATLG		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 23B.

Table 23B. Comparison of NOV23a against NOV23b.		
Protein Sequence	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV23b	1..267	208/267 (77%)
	1..267	208/267 (77%)

Further analysis of the NOV23a protein yielded the following properties shown in Table 23C.

Table 23C. Protein Sequence Properties NOV23a	
PSort analysis:	0.8000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 51 and 52

A search of the NOV23a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 23D.

Table 23D. Geneseq Results for NOV23a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU01924	Human adenylate cyclase polypeptide - Homo sapiens, 1077 aa. [WO200125448-A1, 12-APR-2001]	1..1068 1..1077	1066/1078 (98%) 1066/1078 (98%)	0.0
AAB02008	Type IV adenylyl cyclase - Homo sapiens, 1064 aa. [US6107076-A, 22-AUG-2000]	1..1066 1..1064	976/1073 (90%) 999/1073 (92%)	0.0
AAB02006	Adenylyl cyclase type II-C2 C2 alpha domain - Homo sapiens, 1090 aa. [US6107076-A, 22-AUG-2000]	12..1060 28..1082	605/1064 (56%) 772/1064 (71%)	0.0
AAR94560	Rat adenylyl cyclase - Rattus sp, 1090 aa. [WO9608260-A1, 21-MAR-1996]	12..1060 28..1082	605/1064 (56%) 772/1064 (71%)	0.0
AAE02938	Human adenylate cyclase 25678 - Homo sapiens, 1086 aa. [WO200144453-A1, 21-JUN-2001]	10..1060 22..1078	605/1066 (56%) 771/1066 (71%)	0.0

In a BLAST search of public sequence databases, the NOV23a protein was found to have homology to the proteins shown in the BLASTP data in Table 23E.

Table 23E. Public BLASTP Results for NOV23a				
Protein Accession Number	Protein/Organism/Length	NOV23a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC37757	SEQUENCE 2 FROM PATENT WO0125448 - Homo sapiens (Human), 1077 aa.	1..1068 1..1077	1066/1078 (98%) 1066/1078 (98%)	0.0
Q91WF3	SIMILAR TO ADENYLYL CYCLASE 4 (ADENYLYL CYCLASE TYPE 4) (EC 4.6.1.1) - Mus musculus (Mouse), 1077 aa.	1..1066 1..1077	988/1078 (91%) 1010/1078 (93%)	0.0
P26770	Adenylate cyclase, type IV (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Rattus norvegicus (Rat), 1064 aa.	1..1066 1..1064	976/1073 (90%) 999/1073 (92%)	0.0
P26769	Adenylate cyclase, type II (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Rattus norvegicus (Rat), 1090 aa.	12..1060 28..1082	605/1064 (56%) 772/1064 (71%)	0.0
Q08462	Adenylate cyclase, type II (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Homo sapiens (Human), 887 aa (fragment).	188..1060 1..879	526/888 (59%) 658/888 (73%)	0.0

PFam analysis predicts that the NOV23a protein contains the domains shown in the Table 23F.

Table 23F. Domain Analysis of NOV23a			
Pfam Domain	NOV23a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Sodcu	376..383	6/8 (75%) 8/8 (100%)	0.73
guanylate_cyc	264..448	71/226 (31%) 156/226 (69%)	4.2e-66
guanylate_cyc	855..1055	94/228 (41%) 177/228 (78%)	2.7e-85

Example 24.

5

The NOV24 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 24A.

Table 24A. NOV24 Sequence Analysis			
	SEQ ID NO: 69	997 bp	
NOV24a, CG59955-01 DNA Sequence	TTCTAATCTGTGGTTTCTTCACATCAACTGAAACAATGCAGCAAAATAACAGTGTGAC TGAATTCATACTGTTAGGATTAACACAGGATCCCTTGAGGCAGAAAATAGTGTTTGTA ATCTTCTTAATTTTCTATATGGGAACGTGGTGGGGAATATGCTCATTATTGTGACCA TCAAGTCCAGCCGGACACTAGGAAGCCCCATGTACTTCTTTCTATTTTATTGTCTT TGCAGATTCTTGCTTTTCAACTTCCACAGCCCCTAGATTAATTGTGGATGCTCTCTCT GAAAAGAAAATTATAACCTACAATGAGTGCATGACACAAGTCTTTGCACTACATTAT TGGCTGCATGGAGATCTTTGTCTCATTCTCATGGCTGTTGATCGCTATGTGGCCAT CTGTAAGCCCTTGCGTTACCCAACCATCATGAGCCAGCAGGTCTGCATCATCCTGATT GTTCTTGCCCTGGATAGGGTCTTTAATACACTCTACAGCTCAGATTATCCTGGCCCTAA GATTGCGCTTCTGTGGACCCTATTGATTGATCATTATTGCTGTGATTGTCAGCCCTT GTTGAAACTTGCCCTGCATGGACACTTACATGATCAACCTGCTGTTGGTGTCTAACAGT GGGGCAATTTGCTCAAGTAGTTTCATGATTTTGATAATTCATATATTGTCATCTTGC ATTCAGTGAGAAACACAGTGCCAAAGGGAAGAAAAAGGCTCTCTCCGCTTGCACGCT TCACATAATTGTAGTCATCTTATTCTTTGGCCCATGTATATTATATATACACGCCCC CCGACCACCTTTCCCATGGACAAGATGGTGGCAGTATTTTATACTATTGGACCCCTT TTCTCAATCCACTCATCTACACACTGAGGAATGCAGAAGTAAAAATGCCATGAGAAA GTTATGGCATGGCAAAATTATTTAGAAAACAAAGGATAAATTGAGGGCCTGACCTGA TTACTTTTTCA		
	ORF Start: ATG at 36	ORF Stop: TAA at 966	
	SEQ ID NO: 70	310 aa	MW at 35002.7kD
NOV24a, CG59955-01 Protein Sequence	MQQNNSVTEFILLGLTQDPLRQKIVFVIFLI FYMGTVVGNMLIIVTIKSSRTLGSMPY FFLFYLSFADSCFSTSTAPRLIVDALSEKKIITYNECMQVVALHLFGCMEIFVLILM AVDRYVAICKPLRYPTIMSQQVCIIILIVLAWIGSLIHSTAQIILALRLPFCGPYLIDH YCCDLQPLLKLACMDTYMINLLVSNLGAICSSSFMLIISYIVILHSLRNHSAKGGK KALSACTSHIIVVILFFGPCIIFYTRPPTTFPMDKMVAVFYITIGPPFLNPLIYTLRNA EVKNAMRKLWHGKIISENKG		

Further analysis of the NOV24a protein yielded the following properties shown in Table 24B.

Table 24B. Protein Sequence Properties NOV24a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.0300 probability located in mitochondrial inner membrane
SignalP analysis:	Cleavage site between residues 40 and 41

A search of the NOV24a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 24C.

Table 24C. Geneseq Results for NOV24a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV24a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG72121	Human olfactory receptor polypeptide, SEQ ID NO: 1802 - Homo sapiens, 299 aa. [WO200127158-A2, 19-APR-2001]	1..298 2..299	296/298 (99%) 296/298 (99%)	e-170
AAU24690	Human olfactory receptor AOLFR189 - Homo sapiens, 298 aa. [WO200168805-A2, 20-SEP-2001]	1..286 1..286	284/286 (99%) 284/286 (99%)	e-163
AAE09969	G-protein coupled-receptor (GPCR) NOV5 protein - Unidentified, 310 aa. [WO200166746-A2, 13-SEP-2001]	1..309 1..309	210/309 (67%) 262/309 (83%)	e-125
AAG72412	Human OR-like polypeptide query sequence, SEQ ID NO: 2093 - Homo sapiens, 310 aa. [WO200127158-A2, 19-APR-2001]	1..309 1..309	210/309 (67%) 262/309 (83%)	e-125
AAG72314	Human olfactory receptor polypeptide, SEQ ID NO: 1995 - Homo sapiens, 310 aa. [WO200127158-A2, 19-APR-2001]	1..309 1..309	210/309 (67%) 262/309 (83%)	e-125

In a BLAST search of public sequence databases, the NOV24a protein was found to have homology to the proteins shown in the BLASTP data in Table 24D.

Table 24D. Public BLASTP Results for NOV24a				
Protein Accession Number	Protein/Organism/Length	NOV24a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8VGF4	OLFACTORY RECEPTOR MOR230-3 - Mus musculus (Mouse), 307 aa.	1..307 1..307	272/307 (88%) 294/307 (95%)	e-160
Q8VGF5	OLFACTORY RECEPTOR MOR230-2 - Mus musculus (Mouse), 307 aa.	1..305 1..305	261/305 (85%) 285/305 (92%)	e-155
Q8VGF6	OLFACTORY RECEPTOR MOR230-1 - Mus musculus (Mouse), 307 aa.	1..307 1..307	259/307 (84%) 286/307 (92%)	e-154
Q8VFF8	OLFACTORY RECEPTOR MOR230-6 - Mus musculus (Mouse), 303 aa.	4..304 3..303	238/301 (79%) 273/301 (90%)	e-143
Q8VFF9	OLFACTORY RECEPTOR MOR230-5 - Mus musculus (Mouse), 303 aa.	4..304 3..303	237/301 (78%) 275/301 (90%)	e-141

PFam analysis predicts that the NOV24a protein contains the domains shown in the Table 24E.

Table 24E. Domain Analysis of NOV24a			
Pfam Domain	NOV24a Match Region	Identities/ Similarities for the Matched Region	Expect Value
7tm_1	39..199	36/181 (20%) 120/181 (66%)	4.4e-27

Example 25.

- 5 The NOV25 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 25A.

Table 25A. NOV25 Sequence Analysis			
	SEQ ID NO: 71	847 bp	
NOV25a, CG89211-01 DNA Sequence	<u>AACAATGCAGTGCAGGATGGCATACCATGGCAACAGGGGCAC</u> <u>TTCTCACCAGCCACA</u> TTTTTTCTCATTGGAATCCCAGGTCTGGAAGACGTCCATATGAGAATCTCCCTGCATT TCTGCTCTGTTTACCTTTTGGCTTTGCTGGGAAATGCTACCATTCTGTTAGTCATCAA GGCAGAACAGACCCTCCGGGAGCCCATGTTCTACTTTCTGGCCATCCTTTCCACAATT GATTTGGCCCTTTCTACAACCTCTGTGCCTCGTACGCTGGGTATCTTCTGCGTTTGATG TTCATGAGATTAACCTTTGGAGCTTGTGTGGCCAGATGTTTCTGATCCATGCCCTTCAG TGGCATGGAGGCTGAGGTCTGGTGGCCATGGCCTTTGACCATTACGTGGCCATCTGCA AATCCACTTCACTACACAAACATCTTGACATGCCGGGTGCTGGTGGGCATCACTATGT GCATTTTAATTCTGCCAGTTCTGTTTACACTCCCGATAATCTATCTCATCTACCGTTT ACCATTTTGTGCAGGCTCATATAATAGCCCATTCCTACTATGAGCACATGGGCATTGCA AAATTGTCCTGTGGAACATCCGTGTCAATGCTATCTATGGGCTCTTTGTGGTCTCCG TCTTTCTCCTGAACCTGGTCCCTATTGTTATCTCATATGTTTACATTCTCTGTGCTGT CTTCTGCCTCCCATCATGATGCTCGGCTAAAAGCCCTAAGCACATGTGGCTCTCAT GTTGGGGTCACTCTGTGTTTCTATATCCCGTCGGTCTTCTCTTCTTCTACTCATTGAT TTGGACACAACATTCCACATTACATTACATTCTT		
	ORF Start: ATG at 5	ORF Stop: TGA at 809	
	SEQ ID NO: 72	268 aa	MW at 30031.5kD
NOV25a, CG89211-01 Protein Sequence	MQCRMAYHGNRGTSHPATFFLIGIPGLEDVHMRISLHFCSVYLLALLGNATILLVIK EQLREPMFYFLAILSTIDLALSTTSVPTLGIFFWFDVHEINFGACVAQMFLIHAFTG MEAEVLVAMAFDHYVAICNPLHYTNILTCRVLVGITMCILIRPVLTFLPIIYLIYRL FCQAHIIAHSYYEHMGIAKLSCGNIRVNAIYGLFVVSFLFLNLVLIVISVYILCAVF CLPSHDARLKALSTCGSHVGVICVFYIPSVFSFLTH		

Further analysis of the NOV25a protein yielded the following properties shown in Table 25B.

Table 25B. Protein Sequence Properties NOV25a	
PSort analysis:	0.6850 probability located in endoplasmic reticulum (membrane); 0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 59 and 60

A search of the NOV25a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 25C.

Table 25C. Geneseq Results for NOV25a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV25a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG71560	Human olfactory receptor polypeptide, SEQ ID NO: 1241 - Homo sapiens, 321 aa. [WO200127158-A2, 19-APR-2001]	5..268 1..264	253/264 (95%) 255/264 (95%)	e-146
ABB44529	Human GPCR5b polypeptide SEQ ID NO 18 - Homo sapiens, 311 aa. [WO200174904-A2, 11-OCT-2001]	5..268 1..263	225/264 (85%) 239/264 (90%)	e-127
ABB44530	Human GPCR5c polypeptide SEQ ID NO 20 - Homo sapiens, 311 aa. [WO200174904-A2, 11-OCT-2001]	5..268 1..263	225/264 (85%) 238/264 (89%)	e-127
AAG71562	Human olfactory receptor polypeptide, SEQ ID NO: 1243 - Homo sapiens, 311 aa. [WO200127158-A2, 19-APR-2001]	5..268 1..263	225/264 (85%) 237/264 (89%)	e-127
AAU24572	Human olfactory receptor. AOLFR62 - Homo sapiens, 311 aa. [WO200168805-A2, 20-SEP-2001]	5..268 1..263	225/264 (85%) 237/264 (89%)	e-127

In a BLAST search of public sequence databases, the NOV25a protein was found to have homology to the proteins shown in the BLASTP data in Table 25D.

Table 25D. Public BLASTP Results for NOV25a				
Protein Accession Number	Protein/Organism/Length	NOV25a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8VGV8	OLFACTORY RECEPTOR MOR32-3 - Mus musculus (Mouse), 317 aa.	15..268 10..263	148/254 (58%) 194/254 (76%)	9e-86
Q8VGX4	OLFACTORY RECEPTOR MOR32-2 - Mus musculus (Mouse), 312 aa.	10..268 5..263	151/259 (58%) 194/259 (74%)	6e-85
Q8VG26	OLFACTORY RECEPTOR MOR32-5 - Mus musculus (Mouse), 313 aa.	8..268 3..263	146/261 (55%) 196/261 (74%)	4e-84
Q8VF06	OLFACTORY RECEPTOR MOR32-11 - Mus musculus (Mouse), 312 aa.	15..268 10..263	144/254 (56%) 191/254 (74%)	2e-83
Q8VG28	OLFACTORY RECEPTOR MOR32-4 - Mus musculus (Mouse), 312 aa.	9..268 4..263	148/260 (56%) 188/260 (71%)	2e-80

PFam analysis predicts that the NOV25a protein contains the domains shown in the Table 25E.

Table 25E. Domain Analysis of NOV25a			
Pfam Domain	NOV25a Match Region	Identities/ Similarities for the Matched Region	Expect Value
7tm_1	48..157	30/112 (27%) 76/112 (68%)	2.2e-09

Example 26.

- 5 The NOV26 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 26A.

Table 26A. NOV26 Sequence Analysis			
	SEQ ID NO: 73	614 bp	
NOV26a, CG90530-02 DNA Sequence	CTTAGTGTGGAGCAGTGAACGTGTGTGGTTCCTTCTACTTGGGGATCATGCAGAGAG CTTCACGCTCTGAAGAGAGAGCTGCACATGTTAGCCACAGAGCCACCCCAAGGCATCA ATGTTGGCAAGATAAAGACCAAATGGATGACCTGCGAGCTCAAATATTAGGTGGAGCC AACACACCTTATGAGAAAGGTGTTTTTAAGCTAGAAGTTATCATTCTCTGAGAGGTAC CATTTGAACCTCCTCAGATCCGATTCTCACTCCAATTTATCATCCAAACATTGATTC TGCTGGAAGGATTTGTCTGGATGTTCTCAAATTGCCACCAAATCCTCAGAATTTAA TATAATAAGCCAGCCTTCTCAAGAATGCCAGACAGTGGACAGAGAAGCATGCAAGAC AGAAACAAAGGCTGATGAGGAAGAGATGCTTGATAATCTACCAGAGGCTGGTGACTC CAGAGTACACAACCTCAACACAGAAAAGGCCAGTCAGCTAGTAGGCATAGAAAAG AAATTTTCATCCTGATGTTTAGGGGACTTGTCTGGTTCATCTTAGTTAATGTGTTCTT TGCCAAGGTATCTAAGTTGCCTACCTTGAATTT		
	ORF Start: ATG at 49	ORF Stop: TAG at 541	
	SEQ ID NO: 74	164 aa	MW at 18958.6kD
NOV26a, CG90530-02 Protein Sequence	MQRASRLKRELHMLATEPPPGITCWQDKDQMDLRAQILGGANTPYEKG VFKLEVIIP ERYPFEPPIRFLTP IYHPNIDSAGRICLDVLKLPKSSSEFKYNKPAFLKNARQWTEK HARQKQKADEEEMLDNLPEAGDSRVHNSTQKRKASQLVGIKKFHPDV		

Further analysis of the NOV26a protein yielded the following properties shown in Table 26B.

Table 26B. Protein Sequence Properties NOV26a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

5 A search of the NOV26a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 26C.

Table 26C. Geneseq Results for NOV26a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV26a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB51233	Human ubiquitin-conjugating enzyme 2 protein sequence SEQ ID NO:7 - Homo sapiens, 197 aa. [CN1268564-A, 04-OCT-2000]	1..164 1..197	164/197 (83%) 164/197 (83%)	2e-89
AAM4145 5	Human polypeptide SEQ ID NO 6386 - Homo sapiens, 207 aa. [WO200153312-A1, 26-JUL-2001]	1..164 11..207	164/197 (83%) 164/197 (83%)	2e-89
AAM3966 9	Human polypeptide SEQ ID NO 2814 - Homo sapiens, 197 aa. [WO200153312-A1, 26-JUL-2001]	1..164 1..197	164/197 (83%) 164/197 (83%)	2e-89
AAU23245	Novel human enzyme polypeptide #331 - Homo sapiens, 207 aa. [WO200155301-A2, 02-AUG-2001]	1..164 11..207	164/197 (83%) 164/197 (83%)	2e-89
AAY23157	Human ubiquitin-like conjugating protein (UBCLE) - Homo sapiens, 197 aa. [WO9931252-A1, 24-JUN-1999]	1..164 1..197	164/197 (83%) 164/197 (83%)	2e-89

In a BLAST search of public sequence databases, the NOV26a protein was found to have homology to the proteins shown in the BLASTP data in Table 26D.

Table 26D. Public BLASTP Results for NOV26a				
Protein Accession Number	Protein/Organism/Length	NOV26a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9NPD8	UBIQUITIN-CONJUGATING ENZYME ISOLOG (HSPC150) (UBIQUITIN-CONJUGATING ENZYME E2) (CDNA FLJ20497 FIS, CLONE KAT08890) (HSPC150 PROTEIN SIMILAR TO UBIQUITIN-CONJUGATING ENZYME) - Homo sapiens (Human), 197 aa.	1..164 1..197	164/197 (83%) 164/197 (83%)	4e-89
Q9CQ37	2700084L22RIK PROTEIN - Mus musculus (Mouse), 204 aa.	1..164 1..196	136/197 (69%) 145/197 (73%)	7e-70
Q9XHP3	UBIQUITIN-CONJUGATING ENZYME E2 - Catharanthus roseus (Rosy,periwinkle) (Madagascar periwinkle), 153 aa.	6..91 9..94	49/86 (56%) 60/86 (68%)	2e-23
Q8W011	UBIQUITIN-CONJUGATING ENZYME E2 - Oryza sativa (Rice), 153 aa.	6..91 9..94	46/86 (53%) 58/86 (66%)	7e-21
Q94A97	AT1G78870/F9K20_8 - Arabidopsis thaliana (Mouse-ear cress), 153 aa.	6..91 9..94	46/86 (53%) 57/86 (65%)	7e-21

PFam analysis predicts that the NOV26a protein contains the domains shown in the Table 26E.

Table 26E. Domain Analysis of NOV26a			
Pfam Domain	NOV26a Match Region	Identities/ Similarities for the Matched Region	Expect Value
UQ_con	1..116	51/165 (31%) 91/165 (55%)	2.9e-23

Example 27.

- 5 The NOV27 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 27A.

Table 27A. NOV27 Sequence Analysis			
	SEQ ID NO: 75	973 bp	
NOV27a, CG93076-01 DNA Sequence	GGATGAACCAGACTTTGAATAGCAGTGGGACCGTGGAGTCAGCCCTAAACTATTCCAG AGGGAGCACAGTGCACACGGCCTACCTGGTGTGAGCTCCCTGGCCATGTTACCTGTC CTGTGCGGGATGGCAGGCAACAGCATGGTGATCTGGCTGTCTGGGCTTTCGAATGCAC GGAACCCCTTCTGCATCTATATCCTCAACCTGGCGGCAGCCGACCTCCTTCTCTCT CAGCATGGCTTCCACGCTCAGCCTGGAAACCCAGCCCTGGTCAATACCACTGACAAG GTCCACGAGCTGATGAAGAGACTGATGTACTTTGCCTACACAGTGGGCCTGAGCCTGC TGACGGCCATCAGCACCCAGCGCTGTCTCTGTCTCTTCCCTATCTGGTTCAAGTG TCACCGGCCAGGCACCTGTGAGCCTGGGTGTGTGGCCTGTGTGGACACTCTGTCTC CTGATGAACGGGTTGACCTCTTCTTCTGAGCAAGTTCTTGAATCAATGAAGATC GGTGCTTCAGGGTGGACATGGTCCAGGCCGCCCTCATCATGGGGGTCTTAACCCAGT GATGACTCTGTCCAGCCTGACCCTCTTTGTCTGGGTGCGGAGGGGCTCCAGCAGTGG CGGCGGCAGCCCAACGGCTGTTCTGTGGTGGTCTTGGCCTCTGTCTGGTGTCTCTCA TCTGTTCCCTGCCTCTGAGCATCTACTGGTTTGTGCTCTACTGGTTGAGCCTGCCGCC CGAGATGCAGGTCCTGTGCTTCAGCTTGTACGCCTCTCTCTCGTCCGTAAGCAGCAGC GCCAACCCCGTCATCTACTTCTGTGGGCGCCGAGGAGGCCACAGGCTGCCACCA GGTCCCTGGGGAAGTGTGCTCAACAGGCGCTTCGCGAGGAGCCGAGCTGGAAGGTGG GGAGACGCCCAACCGTGGGCACCAATGAGATGGGGGCTTGAGAGCC		
	ORF Start: ATG at 3	ORF Stop: TGA at 966	
	SEQ ID NO: 76	321 aa	MW at 36087.1kD
NOV27a, CG93076-01 Protein Sequence	MNQTLNSSGTVESALNYSRGSSTVHTAYLVLSLAMFTCLCGMAGNSMVIWLLGFRMHR NPFICIYILNLAADLLFLFSMASTLSLETQPLVNTTDKVHELMKRLMYFAYTVGLSLL TAISTQRCLSVLFPFIWFKCHRPHLSAWVCGLLWTLCLLMNGLTSSFCSKFLKFNEDR CFRVDMVQAALIMGVLTTPVMTLSSLTLFVWVRRGSQQWRRQPTRLFVVVLASVLVFLI CSLPLSIYWFVLYWLSLPPMQVLCFSLSRLLSSSVSSSANPVYFVLGSRSSRLRPLTR SLGTVLQQALREEPELEGGETPTVGTNEMGA		

Further analysis of the NOV27a protein yielded the following properties shown in Table 27B.

Table 27B. Protein Sequence Properties NOV27a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 45 and 46

A search of the NOV27a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 27C.

Table 27C. Geneseq Results for NOV27a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV27a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB04658	Human G protein-coupled receptor TGR7 SEQ ID NO:1 - Homo sapiens, 321 aa. [WO200183748-A1, 08-NOV-2001]	1..321 1..321	320/321 (99%) 320/321 (99%)	0.0
AAU19292	Human G protein-coupled receptor nGPCR-74 - Homo sapiens, 321 aa. [WO200166750-A2, 13-SEP-2001]	1..321 1..321	320/321 (99%) 320/321 (99%)	0.0
AAG64124	Human G protein-coupled receptor GPRv51 - Homo sapiens, 321 aa. [WO200148188-A1, 05-JUL-2001]	1..321 1..321	320/321 (99%) 320/321 (99%)	0.0
AAU04366	Human G-protein coupled receptor, hRUP12 - Homo sapiens, 321 aa. [WO200136471-A2, 25-MAY-2001]	1..321 1..321	320/321 (99%) 320/321 (99%)	0.0
AAE06768	Human G-protein coupled receptor-18 (GCREC-18) protein - Homo sapiens, 321 aa. [WO200157085-A2, 09-AUG- 2001]	1..321 1..321	320/321 (99%) 320/321 (99%)	0.0

In a BLAST search of public sequence databases, the NOV27a protein was found to have homology to the proteins shown in the BLASTP data in Table 27D.

Table 27D. Public BLASTP Results for NOV27a				
Protein Accession Number	Protein/Organism/Length	NOV27a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q91ZB8	G PROTEIN-COUPLED RECEPTOR - Mus musculus (Mouse), 321 aa.	1..318 1..319	189/321 (58%) 237/321 (72%)	3e-98
Q96LA9	G PROTEIN-COUPLED RECEPTOR - Homo sapiens (Human), 322 aa.	1..310 1..305	134/317 (42%) 189/317 (59%)	7e-54
AAL86883	G PROTEIN-COUPLED RECEPTOR SNSR6 - Homo sapiens (Human), 322 aa.	1..310 1..305	133/317 (41%) 188/317 (58%)	2e-53
AAL86882	G PROTEIN-COUPLED RECEPTOR SNSR5 - Homo sapiens (Human), 322 aa.	1..310 1..305	134/317 (42%) 188/317 (59%)	3e-53
AAL86880	G PROTEIN-COUPLED RECEPTOR SNSR3 - Homo sapiens (Human), 322 aa.	1..309 1..304	129/314 (41%) 188/314 (59%)	2e-52

PFam analysis predicts that the NOV27a protein contains the domains shown in the Table 27E.

Table 27E. Domain Analysis of NOV27a			
Pfam Domain	NOV27a Match Region	Identities/ Similarities for the Matched Region	Expect Value
7tm_1	44..78	17/37 (46%) 33/37 (89%)	1.4e-09
7tm_1	104..276	39/205 (19%) 119/205 (58%)	7.3e-09

Example 28.

- 5 The NOV28 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 28A.

Table 28A. NOV28 Sequence Analysis

Table 28A. NOV28 Sequence Analysis			
	SEQ ID NO: 77	2856 bp	
NOV28a, CG94235-01 DNA Sequence	GGGCGGCGCGGGGTCTGCGCTGGGGCCATGGCTCCGCGCGCCGCTTCGTCCTGGAGC TTCCCGACTGCACCTTGGCTCACTTCGCCCTAGGCGCGACGCCCCGGGCGACGCAG CGCCCCGACCCCGCCTGGCGGCGCTGCTGGGGCCCCCGGAGCGCAGCTACTCGCTG TGCGTGCCCGTGACCCCGGACGCCGGCTGCGGGGCCCGGGTCCGGGCGCGGCGCTG ACCAGCGCCTGCTGCACCACTGCGCCGCGGGCCCCCTTCCAGCGGTGCCAGCTGCTCAG GCTGCTCTGCTACTGCCCCGGGCGGCCAGGCCGCGCGCGCACAGCAAGGCTTCCTGCTG CGCGACCCCTGGATGACCTTGACACCCGGCAAGCGCTGCTCGAGCTGCTGGGCGCC GTCAGGAGGCCACCGCCGCACTTGGGCGAGTTCGAGGCCGACCCGCGCGGCCAGCT GTGGCAGCGCCTCTGGGAGGTGCAAGACGGCAGGCGGCTGCAGGTGGGCTGCGCACAG GTCGTGCCCGTCCCGGAGCCCCGCTGCACCCGGTGGTGCCAGACTTGCCAGATTCCG TGGTCTTCCCGGACCGGAAGCCGCCGGGCGGTTTGGAGGAGTGACCTCCTTTAT TCCTGAAGCCCGGGCAGTGCTTGACCTGGTCGACCAGTGCCCAAAACAGATCCAGAA GGAAAGTTCAGGTTGTTGCCATCGAAGGACTGGATGCCAGGGTGGTAAACACCGG TGACCCAGTCAGTGGCAGATTCACTTAAGGCTGTCTCTTAAAGTCACCACTCTTCT CATTTGGCCAGTGGAGGAAGATCTTTGATGATGAACCAACTATCATTAGAAGAGCTTT TACTCTTTGGGCAATTATATTGTGGCCTCCGAAATAGCTAAAGAATCTGCCAAATCTC CTGTGATTGTAGACAGGCACAGCACGGCCACCTATGCCATAGCCACTGAGGTGAGTGG GGGTCTCCAGCACCTGCCCCAGCCCATCACCTGTGTACCACTGGCCAGAGGACCTG CTCAAACCTGACCTTATCCTGCTGCTCACTGTGAGTCTCTGAGGAGAGGTTGCGAGAGG TGCAGGGCCGGGGCATGGAGAAGACCAGGGAAGAAGCAGAACTTGAGGCCAACAGTG GTTTCGTCAAAAGGTAGAAATGTCTTACCAGCGGATGGAGAATCTGGCTGCCATGTG GTTGATGCCAGCCCTCCAGAGAAAAGTCTCTGCAGACGGTATTAAAGCCTAATCCAG ATAGTTTTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGCTCTAAGTATAGATG TTGTTTGAACATCTACATCCACCTTGTGTTATGCAGTGTTCCTCAAATTTCTGTTCTA CAAGCATGTTGTGTGGCAGAAAAGTGGAGACCAGGCATCTTAATTTTACTTCAGCCAT CGTACCCTCTTCTGACTGATGGACCCGTATCAAAAGGTCCTCTCATGTTTCCG GTGAGAGGCCAGCGATTGCTTTCTTCTGGCATAGTAAACATTTTCTTGAACATATG TTTCACTTAATCACTACCAATATCTGGAAGACCTGTCTTACTCAGACAGCAGGAGT GTACAGAAGCAGCAGACAAGATCTTCCAGATCAGCAGGGAGACCCCGGAGCCTCTGCT TCTCTACACTGGCATGCTGATGAGATCGTGACATGCCACATTTGGCTTCTTCCACAT CTGGTTGCACTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAATTTCTAGAG CAAGTGTTCTCTGCAGAGGCTGTCTATGTGTCTTGGCTGCCCAAGGACACTCTGCGAG GCCATTTTGGGTAAGGAACACTTACAAGAAGGCATTGATCTTGTGCTGAGGCTCA GAGCCCTTTTGATAGGCTTCTGAGTCATATATAAAGACATTCAAGCCAAGATGCTCC ACTGCAAAATATACCAACCTTCTCTGAATTATATTTTGGCTTATTTATATTTCTTTCT TTTTCTDPAAGTATGGCTCTGAATAGAATGCACATTTTCCATTGAAGTGGATGCATT CATTTAGCCAATCCAGTAATTTATTTATATTAATCTATACATAATATGTTTCTCCAG ATAGGAGCTATGATTCTTAATTAAGTGGAGTCAAAACGCTAAATGCAATGTTTGT TGTGATTTTCTATTACAAAACCTTAATTTGTCTTGTAAATAAGTACAGTGGATCTTG GAGTGGGATTTCTTGGTAAATATCTTGCACTGAATGTCTCATGATTACATATGAAG TCGCTTTGACATATCTTTAGACAGAAAAAGTAGCTGAGTGAGGGGGAAATTATAGAG CTGTGTGACTTTAGGGAGTAGGTTGAACCAGGTGATTACCTAAATTCCTTCCAGTTC AAAGGCAGATAAATCTGTAATTTATTTATCTATCTACCATTTCTTAAGAAGACATT ACTCCAAAATAATTAATTTAAGGCTTATCAGGTCTGCATATAGAATCTTAAATCT AATAAAGTTTCATGTTAATGTATAGGATTTTAAAGAGCTATAGGTAATTTCTAT TAATATGTGTATATTAATGTAATTGATTTCAAGTTGAAAGTATTTAAAGCTGATA ATAGCATTAGGTTCTTGTCAATGTGGTATCTAGCTGATTATGTTTATTATTCT TAAACATTTTGAAGCTTATACTGGCAGCCTAGAAAAACAAATTAATGTATCT TATGTCCCTGGCACATGAATAAATTTGCTGTGGTTTACTAATCTAAAAAATAAAAA AAAGGGCGGCCGCT		
	ORF Start: ATG at 28	ORF Stop: TAG at 1294	
	SEQ ID NO: 78	422 aa	MW at 46476.6kD
NOV28a, CG94235-01 Protein Sequence	MAPRRFVLELPDCTLAHFALGADAPGDADAPDRLAALLGPPERSYSLCVPVTPDAG CGARVRAARLHQRLHLQRRGPFQRCQLRLCYCPGGQAGGAQGGFLRLDPLDDPD RQALLELLGACQEAAPRHLGEFEADPRGQLWQRLWEVDGRRLLQVGCQVVPPEPPL HPVVPDLPSVVFPDREAARVLEBCTSFIPPEARVLDLVDQCPKQIQKGFQVVAIE GLDATGGKTTVTQSVADSLKAVLLKSPSPSCIGQWRKIFDDEPTIIRRAFYSLGNIVIA SEIAKESAKSPVIVDRHSTATYAIATEVSGGLQHLPPAHFVYQWPEDLLKPDILLLL TVSPEERLQRLQGRGMEKTRREEALEANSVFRQKVMESYQRMENPGCHVVDAASPSEK VLQTVLSLQNSFSEP		

	SEQ ID NO: 79	2331 bp
NOV28b, CG94235-02 DNA Sequence	GTCTGCGCTGGGGCCATGGCTCCGCCGCGCCGCTTCGTCTGGAGCTTCCTGACTGCA CCCTGGCTCACCTCGCCCTAGGCGCCGTTTTGGAGGAGTGTACCTCCTTTATTCCTGA AGCCCGGGCAGTGCTTGACCTGGTCGACCAAGACAGATCCAGAAAGGAAAG TTCCAGGTTGTTGCCATCGAAGGACTGGATGCCACGGGTTAAACACAGGTGACCCAGT CAGCGGCAGATTCACTTAAGGCTGTCCTCTTAAAGTCAACACCTCTTGCATTGGCCA GTGGAGGAAGATCTTTGATGATGAACCACTATCATTAGAAGAGCTTTTACTCTTTG GGCAATTATATTGTGGCCTCCGAAATAGCTAAAGAACTGCCAAATCTCCTGTGATTG TAGACAGGTACTGGACAGCAGGCCACCTATGCCATAGCCACTGAGGTGAGTGGGGG TCTCCAGCACCTGCCCCAGCCCATCACCTGTGTACAGTGGCCAGAGGACCTGCTC AAACCTGACCTTATCTGCTGCTCACTGTGAGTCTGAGGAGAGGTTGCAGAGGCTGC AGGGCCGGGGCATGGAGAAGACCAGGGAAGAAGCAGAACTTGAGGCCAACAGTGTGTT TCGTCAAAAGGTAGAAATGTCCTACCAGCGGATGGAGAATCCTGGCTGCCATGTGGTT GATGCCAGCCCTCCAGAGAAAAGGTCTGCAGACGGTATTAAAGCCTAATCCAGAATA GTTTTAGTGAACCGTAGTTACTCTGGCCAGGTGCCACGTCTAACTAGATTAGATTGTG TTTGAACATCTACATCCACCATTTGTTATGCAGTGTTCCTAAATTTCTGTTCTACAA GCATGTTGTGTGGCAGAAAACCTGGAGACCAGGCATCTTAAGTTTACTTCAGCCATCGT ACCCTCTTCTGACTGATGGACCCGTCATCACAAGGTCCCTCTCATATGTTCCAGTG AGAGGCCAGCGATTGCTTTCTTCTGGCATAGTAAACATTTCTTGAACATATGTTT CACTTAATCACTACCAATATCTGGAAGACCTGTCTTACTCAGACAGCACCAGGTGTA CAGAAGCAGCAGACAAGATCTTCCAGATCAGCAGGGAGACCCGGAGCCTGTCTTCT CCTACACTGGCATGCTGATGAGATCGTGACATGCCACATTGGCTTCTTCCACATCTG GTTGCACTCGTCATGATGGGCTCGCTGCATCTCCCTCAGTCCCAAATTTCTAGAGCCAA GTGTTCTGTCAGAGGCTGTCTATGTGTCCTGGCTGCCCAAGGACACTCTGCGAGAGCC ATTTTGGGTAAGGAACACTTACAAGAAGGCATTGATCTTGTGTCGAGGCTCAGAG CCCTTTTGATAGGCTTCTGAGTCATATATAAAGACATTCAAGCCAAGATGCTCCAAT GCAAATATACCAACCTTCTCTGAATTATATTTGCTTATTTATATTTCTTTCTTTT TTCTAAAGTATGGCTCTGAATAGAATGCACATTTCCATTGAACCTGGATGCATTTAT TTAGCCAATCCAGTAATTTATTTATATTAATCTATACATAATATGTTTCTCAGCATA GGAGCTATGATTCATTAATTAAGTGGAGTCAAAACGCTAAATGCAATGTTTGTGTG GTATTTTCATTACACAACTTAATTTGCTTGTGTTAAATAAGTACAGTGGATCTTGGAG TGGGATTTCTTGGTAAATTTATCTTGCACTTGAATGTCTCATGATTACATATGAATCG CTTTGACATATCTTTAGACAGAAAAAGTAGCTGAGTGAGGGGGAAATATAGAGCTG TGTGACTTTAGGGAGTAGGTTGAACAGGTGATTACCTAAATTCCTTCCAGTTCAAA GGCAGATAAATCTGTAATTTATTTATCTATCTACCATTTCTTAAGAAGACATTACT CCAAATAATTAATTTAAGGCTTTATCAGGTCTGCATATAGAATCTTAAATTTAAT AAAGTTTCATGTTAATGTATAGGATTTTAAAGAGCTATAGGTAATTTCTGTATAA TATGTGTATATTAATGTAAATGATTTGATTCAGTTGAAAGTATTTAAAGCTGATAATA GCATTAGGGTTCTTTGCAATGTGGTATCTAGCTGATTATTTGTTTATTTACTTTAA ACATTTTGAAAGCTTATACTGGCAGCTAGAAAAACAACAATTAATGTATCTTTAT GTCCCTGGCACATGAATAAACTTTGCTGTGTTTACTAAAAAATAAAAAAAAAA GGGCGGCCGCT	
	ORF Start: ATG at 16	ORF Stop: TAG at 769
	SEQ ID NO: 80	251 aa MW at 27980.8kD
NOV28b, CG94235-02 Protein Sequence	MAPPRFVLELPDCTLAHFALGAVLEECTSFIPEARAVLDLVDQCPKQIQKGFQVVA IEGLDATGKTTVTQSAADSLKAVLLKSPSCIGQWRKIFDDEPTIIRAFYSLGNIV ASEIAKESAKSPVIVDRYWHSTATYAIATEVSGGLQHLPPAHPVYQWPEDLLKPDLI LLLTVSPEERLQRLQGRGMEKTRREEALEANSVFRQKVMSYQRMENPGCHVVDASPS REKVLQTVLSLIQNSFSEP	

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 28B.

Table 28B. Comparison of NOV28a against NOV28b.		
Protein Sequence	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV28b	192..422 20..251	217/233 (93%) 217/233 (93%)

Further analysis of the NOV28a protein yielded the following properties shown in Table 28C.

Table 28C. Protein Sequence Properties NOV28a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1939 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV28a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 28D.

Table 28D. Geneseq Results for NOV28a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB57297	Mouse ischaemic condition related protein sequence SEQ ID NO:834 - Mus musculus, 431 aa. [WO200188188-A2, 22-NOV-2001]	199..384 157..344	131/193 (67%) 150/193 (76%)	3e-64
AAU34536	E. coli cellular proliferation protein #117 - Escherichia coli, 213 aa. [WO200170955-A2, 27-SEP-2001]	229..421 6..210	51/212 (24%) 88/212 (41%)	0.054
AAB72201	E. coli thymidylate kinase amino acid sequence - Escherichia coli, 213 aa. [WO200111025-A2, 15-FEB-2001]	229..421 6..210	51/212 (24%) 88/212 (41%)	0.054
AAU34536	E. coli cellular proliferation protein #117 - Escherichia coli, 213 aa. [WO200170955-A2, 27-SEP-2001]	229..421 6..210	51/212 (24%) 88/212 (41%)	0.054
AAU28786	E.coli thymidylate kinase-I - Escherichia coli, 213 aa. [WO9941404-A2, 19-AUG-1999]	229..421 6..210	51/212 (24%) 88/212 (41%)	0.054

In a BLAST search of public sequence databases, the NOV28a protein was found to have homology to the proteins shown in the BLASTP data in Table 28E.

Table 28E. Public BLASTP Results for NOV28a				
Protein Accession Number	Protein/Organism/Length	NOV28a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9DC34	I200004E04RIK PROTEIN - Mus musculus (Mouse), 395 aa.	42..418 16..393	292/379 (77%) 328/379 (86%)	e-168
Q62316	THYMIDYLATE KINASE HOMOLOGUE - Mus musculus (Mouse), 431 aa.	199..384 157..344	131/193 (67%) 150/193 (76%)	7e-64
Q96AL8	HYPOTHETICAL 6.7 KDA PROTEIN - Homo sapiens (Human), 58 aa (fragment).	365..422 1..58	58/58 (100%) 58/58 (100%)	3e-25
O60970	TKRP1 - Leishmania major, 274 aa.	199..418 44..254	73/220 (33%) 108/220 (48%)	9e-22
Q9CKE9	Thymidylate kinase (EC 2.7.4.9) (dTTP kinase) - Pasteurella multocida, 209 aa.	224..416 4..203	59/210 (28%) 87/210 (41%)	4e-04

PFam analysis predicts that the NOV28a protein contains the domains shown in the Table 28F.

Table 28F. Domain Analysis of NOV28a			
Pfam Domain	NOV28a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Thymidylate_kin	231..411	53/208 (25%) 131/208 (63%)	4.9e-08

Example 29.

- 5 The NOV29 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 29A.

Table 29A. NOV29 Sequence Analysis			
	SEQ ID NO: 81	867 bp	
NOV29a, CG94692-01 DNA Sequence	ATGCCGGTGGAGGAATTTGTGGCTGGCTGGATCTCTGGCGCTCTGGGCTTGGTCCTGG GACACCCGTTTGACACTGTAAAGGTGAGGCTGCAGACCCAGACCACCTACCGGGGCAT CGTTGATTGCATGGTCAAGATTTACGCCCATGAGTCCCTCCTGGGCTTCTTCAAGGGA ATGAGCTTCCCCATTGCCAGCATAGCTGTGGTCAACTCTGTCTGTTTGGGGTCTATA GCAACACCCTGCTGGTGCTCACGGCCACCTCCCACCAGGAGCGGCGGGCCAGCCGCC CAGCTACATGCACATCTTCTAGCGGGCTGCACCGGGGGTCTCTGCAGGCCCTACTGT CTGGCTCCTTTTGACCTCATCAAAGTCCGGCTACAAAACCAGACAGAGCCAGGGCCC AGCCAGGGAGCCCCACCCCGGTACCAGGGGGCCGTGCACTGTGCAGCCTCCATCTT CCGGGAGGAGGGGCCCCGGGGGCTGTTCCGAGGAGCCTGGGCCCTGACGCTGAGGGAC ACCCCCACGGTGGGGATCTACTTCATCACCTATGAAGGGCTCTGTGCGCCAGTACACAC CAGAAGGCCAGAATCCAGCTCAGCCACGGTGTCTGGTGGCAGGGGGCTTGCAGGCAT TGCTTCTCTGGGTGGCAGCCACGCCCTTAGACGTGATCAAGTCCCGGATGCAGATGGAT GGACTGAGACGCAGAGTGTACCAGGGGATGCTGGACTGCATGGTGAACAGATCCGGC AGGAAGGACTGGGAGTCTTCTTCCGGGGGTCAACATCAACAGTGCCCGCGCTTTCC CGTCAATGCTGTACCTTCTCAGCTACGAATATCTCCTCCGCTGGTGGGGATGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 865	
	SEQ ID NO: 82	288 aa	MW at 31873.6kD
NOV29a, CG94692-01 Protein Sequence	MPVEEFVAGWISGALGLVLGHPFDTVKVRLQTQTTYRGIVDCMVKIYRHESLLGFFKG MSFPFIASIAVNSVLFGVYSNTLLVLTATSHQERRAQPPSYMHI FLAGCTGGFLQAYC LAPFDLKVRLQNQTEPRAQPGSPPPRYQGPVHCAASI FREEGPRGLFRGAWALTLRD TPTVGIYFITYEGLCRQYTPGQNPSSATVLVAGGFAGIASWVAATPLDVIKSRMQMD GLRRRVYQGMLDCMVSSIRQEGLGVFFRGVTINSARAFPVNAVTFLSYEYLLRWVG		
	SEQ ID NO: 83	867 bp	
NOV29b, CG94692-02 DNA Sequence	ATGCCGGTGGAGGAATTTGTGGCTGGCTGGATCTCTGGCGCTCTGGGCTTGGTCCTGG GACACCCGTTTGACACTGTAAAGGTGAGGCTGCAGACCCAGACCACCTACCGGGGCAT CGTTGATTGCATGGTCAAGATTTACGCCCATGAGTCCCTCCTGGGCTTCTTCAAGGGA ATGAGCTTCCCCATTGCCAGCATAGCTGTGGTCAACTCTGTCTGTTTGGGGTCTATA GCAACACCCTGCTGGTGCTCACGGCCACCTCCCACCAGGAGCGGCGGGCCAGCCGCC CAGCTACATGCACATCTTCTAGCGGGCTGCACCGGGGGTCTCTGCAGGCCCTACTGT CTGGCTCCTTTTGACCTCATCAAAGTCCGGCTACAAAACCAGACAGAGCCAGGGCCC AGCCAGGGAGCCCCACCCCGGTACCAGGGGGCCGTGCACTGTGCAGCCTCCATCTT CCGGGAGGAGGGGCCCCGGGGGCTGTTCCGAGGAGCCTGGGCCCTGACGCTGAGGGAC ACCCCCACGGTGGGGATCTACTTCATCACCTATGAAGGGCTCTGTGCGCCAGTACACAC CAGAAGGCCAGAATCCAGCTCAGCCACGGTGTCTGGTGGCAGGGGGCTTGCAGGCAT TGCTTCTCTGGGTGGCAGCCACGCCCTTAGACGTGATCAAGTCCCGGATGCAGATGGAT GGACTGAGACGCAGAGTGTACCAGGGGATGCTGGACTGCATGGTGAACAGATCCGGC AGGAAGGACTGGGAGTCTTCTTCCGGGGGTCAACATCAACAGTGCCCGCGCTTTCC CGTCAATGCTGTACCTTCTCAGCTACGAATATCTCCTCCGCTGGTGGGGATGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 865	
	SEQ ID NO: 84	288 aa	MW at 31873.6kD
NOV29b, CG94692-02 Protein Sequence	MPVEEFVAGWISGALGLVLGHPFDTVKVRLQTQTTYRGIVDCMVKIYRHESLLGFFKG MSFPFIASIAVNSVLFGVYSNTLLVLTATSHQERRAQPPSYMHI FLAGCTGGFLQAYC LAPFDLKVRLQNQTEPRAQPGSPPPRYQGPVHCAASI FREEGPRGLFRGAWALTLRD TPTVGIYFITYEGLCRQYTPGQNPSSATVLVAGGFAGIASWVAATPLDVIKSRMQMD GLRRRVYQGMLDCMVSSIRQEGLGVFFRGVTINSARAFPVNAVTFLSYEYLLRWVG		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 29B.

Table 29B. Comparison of NOV29a against NOV29b.		
Protein Sequence	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Region

NOV29b	1..288	288/288 (100%)
	1..288	288/288 (100%)

Further analysis of the NOV29a protein yielded the following properties shown in Table 29C.

Table 29C. Protein Sequence Properties NOV29a	
PSort analysis:	0.7900 probability located in plasma membrane; 0.6400 probability located in microbody (peroxisome); 0.3000 probability located in Golgi body; 0.2000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV29a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 29D.

Table 29D. Geneseq Results for NOV29a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE10330	Human transporter and ion channel-7 (TRICH-7) protein - Homo sapiens, 288 aa. [WO200162923-A2, 30-AUG-2001]	1..288 1..288	288/288 (100%) 288/288 (100%)	e-170
ABG27643	Novel human diagnostic protein #27634 - Homo sapiens, 346 aa. [WO200175067-A2, 11-OCT-2001]	1..288 59..346	286/288 (99%) 286/288 (99%)	e-168
ABG27643	Novel human diagnostic protein #27634 - Homo sapiens, 346 aa. [WO200175067-A2, 11-OCT-2001]	1..288 59..346	286/288 (99%) 286/288 (99%)	e-168
AAE16774	Human transporter and ion channel-11 (TRICH-11) protein - Homo sapiens, 181 aa. [WO200192304-A2, 06-DEC-2001]	113..288 6..181	176/176 (100%) 176/176 (100%)	e-101
AAM39422	Human polypeptide SEQ ID NO 2567 - Homo sapiens, 311 aa. [WO200153312-A1, 26-JUL-2001]	3..285 6..300	144/295 (48%) 183/295 (61%)	4e-73

In a BLAST search of public sequence databases, the NOV29a protein was found to have homology to the proteins shown in the BLASTP data in Table 29E.

Table 29E. Public BLASTP Results for NOV29a				
Protein Accession Number	Protein/Organism/Length	NOV29a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8VBZ7	HYPOTHETICAL 32.4 KDA PROTEIN - Mus musculus (Mouse), 294 aa (fragment).	1..287 7..293	246/287 (85%) 267/287 (92%)	e-147
CAC24997	SEQUENCE 41 FROM PATENT WO0100806 PRECURSOR - Homo sapiens (Human), 308 aa.	5..285 2..301	128/301 (42%) 187/301 (61%)	1e-66
Q922X4	SIMILAR TO CG4995 GENE PRODUCT - Mus musculus (Mouse), 306 aa.	5..286 4..275	117/288 (40%) 166/288 (57%)	1e-53
Q9VKZ5	CG4995 PROTEIN (GH13054P) - Drosophila melanogaster (Fruit fly), 399 aa.	3..284 41..309	112/287 (39%) 155/287 (53%)	7e-51
Q9AX03	PUTATIVE CARNITINE/ACYLCARNITINE TRANSLOCASE - Oryza sativa (Rice), 322 aa.	4..286 14..307	119/300 (39%) 166/300 (54%)	3e-50

PFam analysis predicts that the NOV29a protein contains the domains shown in the Table 29F.

Table 29F. Domain Analysis of NOV29a			
Pfam Domain	NOV29a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	1..90	31/125 (25%) 72/125 (58%)	5.1e-22
mito_carr	98..198	34/126 (27%) 84/126 (67%)	7.1e-29
mito_carr	200..288	33/125 (26%) 75/125 (60%)	7.1e-24

Example 30.

The NOV30 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 30A.

Table 30A. NOV30 Sequence Analysis			
	SEQ ID NO: 85	947 bp	
NOV30a, CG94724-01 DNA Sequence	ACTGACCATGGCCGACCAGCCAAAGCCCATCAGCCTGCTCAAGAACCTGCTGGCCAGTGGCTTTGGTGGCATGTGCCTGGTGTTCATGGGGCACCCCTCTGGACACTGTCAAGGTCCGACTGCAGACACAGCCACCGAGTTTGCCTAGACAGCCTCCCATGTACTCTGGGACCTTTGACTCTCTCCCAAAGACTCTTAGAAGGGACATCACAGGGCTATATAAGGGAATGGCTGCCCTATCATCGGTGTCAACCCCATCTTTGCTGTGTGCTTCTTTGGGTTTGGTTTGGGAAGAACTAGGACAGAAACACCCAGAAGATGTGCTCAGCTATCCCAGCTGTTTGCAGCTGGGATGTTATCTGGTACATTCAACACAGGCATCGCGACCCCTGGAGAACCCATAAGTCCTTGTTACACTTTTCAAGCCTTCTTCAGGGGAAACCAAGTACACTGGAACCTTGGACTGTGCAAGAAGTTGTACCAGGAGTTTCAGATCCGAGGCATCTACAAAGGGACCGTGCTTACCCTTATGAGAGACGTTCCAGCTAGCGGAACGTATTTTCATGACAAATGAATGCTAAAAAATATCTTCACTCCGGAAGGAAAGAGGGTCTGTGAGCTCAGTGTGCCTAGAACTCTGGTGGCTGGGTGCATTGCAGGGATCTTCAACTGGGCAATGGCAGTCCCAGCAAGATGTGCTCAAGTATCCCTTCCAGACGGCACCTCCTGGGAAATATCCTAACAGTTTGGGATGTGTGCTGAGGGAGCTGATCTGGGATGAAGGAATCACATCCTTGTCTAAAGGGTCTCATGCAGTGATGACCCGAGCCTTCCAGCCAATGCAGCCTGTTTCTTGGCCTTGAAGTTGCCATGAAGTTTCTTAATTGGGCCACCCCAACTTGTGAGGCAGAGGCTGCTCAAGACTTCTGGATGCTGGAAC		
	ORF Start: ATG at 8	ORF Stop: TGA at 908	
	SEQ ID NO: 86	300 aa	MW at 32855.1kD
NOV30a, CG94724-01 Protein Sequence	MADQPKPISLLKNLLASGFGGMCLVFMGHPLDVTVKVRLQTQPPSLPRQPPMYSGTFDSLPKTLRRDITGLYKGM AAP IIGVTP IFAVCFFGFLRKKLGQKHPEDVLSYPQLFAAGMLSGTFTTG IATPGEPIKSLHLHFQPSSETKYTGTLDCAKKLYQEFQIRGIYKGTVL LMRDVPASGTYFMTNEWLNKNIIFTPEGKRVCELSVPRILVAGCIAGIFNWAMAVPQDVLYKYPFQTAPPKYPNSFGDVLRELIWDEGITSLSKGSDAVMTRAFPANAACFLGLEVAIKFLNWATPNL		

Further analysis of the NOV30a protein yielded the following properties shown in
 5 Table 30B.

Table 30B. Protein Sequence Properties NOV30a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV30a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 30C.

Table 30C. Geneseq Results for NOV30a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV30a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM79434	Human protein SEQ ID NO 3080 - Homo sapiens, 318 aa. [WO200157190-A2, 09-AUG-2001]	1..300 18..318	258/301 (85%) 268/301 (88%)	e-146
AAM78450	Human protein SEQ ID NO 1112 - Homo sapiens, 318 aa. [WO200157190-A2, 09-AUG-2001]	1..300 18..318	258/301 (85%) 268/301 (88%)	e-146
AAY25740	Human secreted protein encoded from gene 30 - Homo sapiens, 228 aa. [WO9938881-A1, 05-AUG-1999]	74..300 1..227	196/227 (86%) 202/227 (88%)	e-110
ABB59582	Drosophila melanogaster polypeptide SEQ ID NO 5538 - Drosophila melanogaster, 306 aa. [WO200171042-A2, 27-SEP-2001]	11..299 16..305	137/291 (47%) 190/291 (65%)	1e-73
ABB59928	Drosophila melanogaster polypeptide SEQ ID NO 6576 - Drosophila melanogaster, 299 aa. [WO200171042-A2, 27-SEP-2001]	11..293 15..297	121/285 (42%) 172/285 (59%)	7e-57

In a BLAST search of public sequence databases, the NOV30a protein was found to have homology to the proteins shown in the BLASTP data in Table 30D.

Table 30D. Public BLASTP Results for NOV30a				
Protein Accession Number	Protein/Organism/Length	NOV30a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
O43772	Mitochondrial carnitine/acylcarnitine carrier protein (Carnitine/acylcarnitine translocase) (CAC) - Homo sapiens (Human), 301 aa.	1..300 1..301	258/301 (85%) 268/301 (88%)	e-146
Q9Z2Z6	MCAC PROTEIN - Mus musculus (Mouse), 301 aa.	1..300 1..301	243/301 (80%) 257/301 (84%)	e-138
P97521	Mitochondrial carnitine/acylcarnitine carrier protein (Carnitine/acylcarnitine translocase) (CAC) - Rattus norvegicus (Rat), 301 aa.	1..300 1..301	233/301 (77%) 254/301 (83%)	e-134
Q9VQG4	COLT PROTEIN (GM13207P) - Drosophila melanogaster (Fruit fly), 306 aa.	11..299 16..305	137/291 (47%) 190/291 (65%)	3e-73
O01396	COLT PROTEIN (CONGESTED-LIKE TRACHEA PROTEIN) - Drosophila melanogaster (Fruit fly), 306 aa.	11..299 16..305	136/291 (46%) 188/291 (63%)	4e-72

PFam analysis predicts that the NOV30a protein contains the domains shown in the Table 30E.

Table 30E. Domain Analysis of NOV30a			
Pfam Domain	NOV30a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	9..104	37/125 (30%) 79/125 (63%)	3.7e-23
mito_carr	108..202	34/125 (27%) 76/125 (61%)	2.4e-19
mito_carr	207..299	28/125 (22%) 70/125 (56%)	1.6e-11

Example 31.

- 5 The NOV31 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 31A.

Table 31A. NOV31 Sequence Analysis

Table 31A. NOV31 Sequence Analysis			
	SEQ ID NO: 87	1171 bp	
NOV31a, CG94871-01 DNA Sequence	GTTGGCTCCAGACAAATAAACATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCAC TTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTAGCCCTGTGGG ATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAAGATGAGAATGGCAGAAGGA GGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAATATGGATG ACAGTGGTTTTTCTCTATTGAGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGA ACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAA AGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACCTTTTCTGGCTCAATTACAACAGGAAGGTTATTCTATATTGTTGTTAAGGGT GATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGGTCCAACAGATGC ATCGACCAAAACTTATTGGAGAAGAATTAGCACAATAAAAGAGCAAAGAGTCCATAA AACAGACCTGGAACGAGTGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGATG GAGGAGGATTGTCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATG AGGAAGCAGATCTCCGACAGGGCTATTGAGTAAAGTATGCAAGGTAGTCCAGAAACAT ATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTCGGAAG AGACGAGAAGCCTACTTTGAAAAACAGCAGCAGCAGCAGGGGACCTATCAGGACAGA GTTCACATCCATGTGAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGG TGATGCTATGAGTGAAGAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTAGAACT GTCAGAAATGATTGAAAAACAGAAAGGAAAAAATAATACCTTTAAAAAATAATTTAGA TATTCATACTTTCCAACATTATCCTGTGTGATTACGATAGGGTCCACTTTGGTAAT GTGTCAAAGAG		
	ORF Start: ATG at 22	ORF Stop: TAA at 1078	
	SEQ ID NO: 88	352 aa	MW at 40068.4kD
NOV31a, CG94871-01 Protein Sequence	MESIFHEKQEGSLCAQHCLNLLQGEYFSPVELSSIAHQLDEEEKMRMAEGGVTSEDY RTFLQPPSGNMDDSGFFSIQVINSALKVWGLELILFNSPEYQRLRIDPINERSFICNY KEHWFTVRKLKGQWFNLNSLLTGPELISDTYLAFLAQLQEGYSIFVVKGLDLPCEA DQLLQMRVQQMHRPKLIGEELAQLKEQVRVHKTDLERVLEANDGSGMLDEDEEDLQRA LALSREQIDMEDEEADLRRAIQLSMQGSSRNISQDMTQTSNTLSEELRKRREAYFE KQQQQQGGDLGSGSSHPERPATSSGALGSDLGDAMSEEDMLQAAVTMSLETVRNDLKT EGKK		
	SEQ ID NO: 89	1219 bp	
NOV31b, CG94871-02 DNA Sequence	GTTGGCTCCAGACAAATAAACATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCAC TTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAATATTTAGCCCTGTGGG ATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGATGAGAATGGCAGAAGGA GGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGCAGCCTTCTGGAATATGGATG ACAGTGGTTTTTCTCTATTGAGTTATAAGCAATGCCTTGAAAGTTTGGGGTTTAGA ACTAATCCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGATCCTATAAATGAA AGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGAAAATTAGGAAAA AGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAATTAATATCAGATACATATCT TGCACCTTTTCTGGCTCAATTACAACAGGAAGGTTATTCTATATTGTTGTTAAGGGT GATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGGGTCCAACAGATGC ATCGACCAAAACTTATTGGAGAAGAATTAGCACAATAAAAGAGCAAAGAGTCCATAA AACCGACCTGGAACGAATGTTAGAAGCAAATGATGGCTCAGGAATGTTAGACGAAGAT GAGGAGGATTGTCAGAGGGCTCTGGCACTAAGTCGCCAAGAAATTGACATGGAAGATG AGGAAGCAGATCTCCGACAGGGCTATTGAGTAAAGTATGCAAGGTAGTCCAGAAACAT ATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGAAGAGCTTCGGAAG AGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGC AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCGGACCTATCAGGACAGAGTTACATCC ATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTGCTATG AGTGAAGAAGACATGCTTCAGGCAGCTGTGACCATGTCTTTAGAACTGTGAGAAATG ATTTGAAAAACAGAAAGGAAAAAATAATACCTTTAAAAAATAATTTAGATATTCATACT TTCCAACATTATCCTGTGTGATTACAGCATAGGGTCCACTTTGGTAATGTGTCAAAGA G		
	ORF Start: ATG at 22	ORF Stop: TAA at 1126	
	SEQ ID NO: 90	368 aa	MW at 42277.8kD
NOV31b, CG94871-02 Protein Sequence	MESIFHEKQEGSLCAQHCLNLLQGEYFSPVELSSIAHQLDEEEKMRMAEGGVTSEDY RTFLQPPSGNMDDSGFFSIQVINSALKVWGLELILFNSPEYQRLRIDPINERSFICNY KEHWFTVRKLKGQWFNLNSLLTGPELISDTYLAFLAQLQEGYSIFVVKGLDLPCEA DQLLQMRVQQMHRPKLIGEELAQLKEQVRVHKTDLERMLEANDGSGMLDEDEEDLQRA LALSREQIDMEDEEADLRRAIQLSMQGSSRNISQDMTQTSNTLSEELRKRREAYFE KQQQKQQQQQQQQQQQQQQQQQQQQQORDLSGQSSHPERPATSSGALGSDLGDAMSEEDMLQ AAVTMSLETVRNDLKTGEGKK		

	SEQ ID NO: 91	1153 bp	
NOV31c, CG94871-04 DNA Sequence	GTTGGCTCCAGACAAATAAACATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCAC TTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAAATATTTAGCCCTGTGGA ATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGA GGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGGTTATAAGCAATGCCTTGAAAG TTTGGGGTTTAGAACTATTCTCTGTTCAACAGTCCAGAGTATCAGAGGCTCAGGATCGA TCCTATAAATGAAAGATCATTTATATGCAATTATAAGGAACACTGGTTTACAGTTAGA AAATTAGGAAAACAGTGGTTTAACTTGAATTCTCTCTTGACGGGTCCAGAAATTAATAT CAGATAGATATCTTGCACTTTTCTTGGCTCAATTACAACAGGAAGGTTATTCTATATT TGTCGTTAAGGGTGATCTGCCAGATTGCGAAGCTGACCAACTCCTGCAGATGATTAGG GTCCAACAGATGCATCGACCAAACTTATTGGAGAAGAATTAGCAACAATAAAAGAGC AAAGAGTCCATAAAACAGACCTGGAACGAGTGTTAGAAGCAAATGATGGCTCAGGAAT GTTAGACGAAGATGAGGAGGATTGCGAGAGGGCTCTGGCACTAAGTCGCCATGAAATT GACATGGAAGATGAGGAAGCAGATCTCCGAGGGCTATTGAGTAAAGTATGCAAGGTA GTTCAGAAACATATCTCAAGATATGACACAGACATCAGGTACAAATCTTACTTCAGA AGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAAACAGCAGCAAAAGCAGCAACAG CAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGACAGAGTTCACATCCATGTGAAA GGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCTAGGTGATGCTATGAGTGAAGA AGACATGCTTCGGGCAGCTGTGACCATGTCTTTAGAAACTGTGCAAAATGATTGAAA ACAGAAGGAAAAAATAATACCTTTAAAAAATAATTTAGATATTCTACTTTCCAACA TTATCCTGTGTGATTACAGCATAGGGTCCACTTTGGTAATGTGTCAAAGAG		
	ORF Start: ATG at 22	ORF Stop: TAA at 1060	
	SEQ ID NO: 92	346 aa	MW at 39764.1kD
NOV31c, CG94871-04 Protein Sequence	MESIFHEKQEGSLCAQHCLNNLLQGEYFSPVELSSIAHQLEDEEERMRMAEGGVTS EDYRTFLQVISNALKVWGLELFLFNSPEYQRLRIDPINERSFCINYKEHWFTVRKLKQWF NLNLLTGPGLISDRYLALFLAQLQQEGYSIFVVKGDLPDCEADQLQMIRVQQMHRP KLIGEELAQKKEQVRVHKTDLERVLEANDGSGMLDEDEEDLQALALSRHEIDMEDEEA DLRRAIQLSMQSSRNISQDMTQTSNTLTSEELRKRREAYFEKQQKQQQQQQQQQ GDLSGQSSHP CERPATSSGALGSDLGDAMSEEDMLRAAVTMSLETVRNLDLKTGKK		
	SEQ ID NO: 93	1000 bp	
NOV31d, CG94871-05 DNA Sequence	GTTGGCTCCAGACAAATAAACATGGAGTCCATCTTCCACGAGAAACAAGAAGGCTCAC TTTGTGCTCAACATTGCCTGAATAACTTATTGCAAGGAGAAATATTTAGCCCTGTGGA ATTATCCTCAATTGCACATCAGCTGGATGAGGAGGAGAGGATGAGAATGGCAGAAGGA GGAGTTACTAGTGAAGATTATCGCACGTTTTTACAGTGGTTTAACTTGAATCTCTCTCT TGACGGGTCCAGAATTAATATCAGATACATATCTTGCACTTTTCTTGGCTCAATTACA ACAGGAAGGTTATTCTATATTGTCTTAAGGGTGATCTGCCAGATTGCGAAGCTGAC CAACTCCTGCAGATGATTAGGGTCCAACAGATGCATCGACCAAACTTATTGGAGAAG AATTAGCACAACTAAAAGAGCAAAGAGTCCATAAAACAGACCTGGAACGAGTGTAGA AGCAAATGATGGCTCAGGAATGTTAGACGAAGATGAGGAGGATTGCGAGGGCTCTG GCACTAAGTCGCCAAGAAATTGACATGGAAGATGAGGAAGCAGATCTCCGCGGGGCTA TTCAGCTAAGTATGCAAGGTAGTTCAGAAACATATCTCAAGATATGACACAGACATC AGGTACAAATCTTACTTCAGAAGAGCTTCGGAAGAGACGAGAAGCCTACTTTGAAAA CAGCAGCAAAAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGGGGGACCTATCAGGAC AGAGTTCACATCCATGTGAAAGGCCAGCCACCAGTTCAGGAGCACTTGGGAGTGATCT AGGTGATGCTATGAGTGAAGAAGACATGCTTCGGGCAGCTGTGACCATGTCTTTAGAA ACTGTGAGAAATGATTGAAAAACAGAAAGAAAAAATAATACCTTTAAAAAATAATTT AGATATTCTACTTTCCAACATTATCTGTGTGATTACAGCATAGGGTCCACTTTGGT AATGTGTCAAAGAG		
	ORF Start: ATG at 22	ORF Stop: TAA at 907	
	SEQ ID NO: 94	295 aa	MW at 33420.8kD
NOV31d, CG94871-05 Protein Sequence	MESIFHEKQEGSLCAQHCLNNLLQGEYFSPVELSSIAHQLEDEEERMRMAEGGVTS EDYRTFLQWFNLNLLTGPGLISDTYLALFLAQLQQEGYSIFVVKGDLPDCEADQLQMIR VQQMHRPKLIGEELAQKKEQVRVHKTDLERVLEANDGSGMLDEDEEDLQALALSRQEI DMEDEEADLRGAIQLSMQSSRNISQDMTQTSNTLTSEELRKRREAYFEKQQKQQQ QQQQQQGDLGSGQSSHP CERPATSSGALGSDLGDAMSEEDMLRAAVTMSLETVRNLDL KTEGKK		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 31B.

Table 31B. Comparison of NOV31a against NOV31b through NOV31d.		
Protein Sequence	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV31b	1..352 1..368	343/368 (93%) 345/368 (93%)
NOV31c	1..352 1..346	326/361 (90%) 328/361 (90%)
NOV31d	129..352 63..295	216/233 (92%) 217/233 (92%)

Further analysis of the NOV31a protein yielded the following properties shown in Table 31C.

Table 31C. Protein Sequence Properties NOV31a	
PSort analysis:	0.3000 probability located in nucleus; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen); 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV31a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 31D.

Table 31D. Geneseq Results for NOV31a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY33492	Human Machado-Joseph disease associated protein - Homo sapiens, 360 aa. [WO9945944-A1, 16-SEP- 1999]	1..326 1..347	321/347 (92%) 323/347 (92%)	0.0
AAR96128	Human Machado-Joseph disease-related protein - Homo sapiens, 291 aa. [JP08092289- A, 09-APR-1996]	1..291 1..291	289/291 (99%) 291/291 (99%)	e-166
ABG22866	Novel human diagnostic protein #22857 - Homo sapiens, 454 aa. [WO200175067-A2, 11-OCT- 2001]	129..291 292..454	163/163 (100%) 163/163 (100%)	2e-87
ABG22866	Novel human diagnostic protein #22857 - Homo sapiens, 454 aa. [WO200175067-A2, 11-OCT- 2001]	129..291 292..454	163/163 (100%) 163/163 (100%)	2e-87
ABG17452	Novel human diagnostic protein #17443 - Homo sapiens, 731 aa. [WO200175067-A2, 11-OCT- 2001]	155..321 16..191	163/176 (92%) 165/176 (93%)	1e-85

In a BLAST search of public sequence databases, the NOV31a protein was found to have homology to the proteins shown in the BLASTP data in Table 31E.

Table 31E. Public BLASTP Results for NOV31a				
Protein Accession Number	Protein/Organism/Length	NOV31a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96TC4	ATAXIN-3 - Homo sapiens (Human), 361 aa.	1..352 1..361	351/361 (97%) 352/361 (97%)	0.0
O15284	JOSEPHIN MJD1 - Homo sapiens (Human), 361 aa.	1..352 1..361	350/361 (96%) 351/361 (96%)	0.0
O15286	JOSEPHIN MJD1 - Homo sapiens (Human), 373 aa (fragment).	2..352 1..373	347/373 (93%) 349/373 (93%)	0.0
Q96TC3	ATAXIN-3 - Homo sapiens (Human), 364 aa.	1..326 1..335	323/335 (96%) 324/335 (96%)	0.0
O15285	JOSEPHIN MJD1 - Homo sapiens (Human), 364 aa.	1..326 1..335	322/335 (96%) 323/335 (96%)	0.0

PFam analysis predicts that the NOV31a protein contains the domains shown in the Table 31F.

Table 31F. Domain Analysis of NOV31a			
Pfam Domain	NOV31a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Josephin	1..198	110/209 (53%) 198/209 (95%)	2.8e-139
UIM	223..240	11/18 (61%) 17/18 (94%)	0.011
UIM	243..260	8/18 (44%) 16/18 (89%)	0.0044
UIM	325..342	9/18 (50%) 17/18 (94%)	0.02

Example 32.

- 5 The NOV32 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 32A.

Table 32A. NOV32 Sequence Analysis

	SEQ ID NO: 95	6224 bp
NOV32a, CG94946-01 DNA Sequence	<p> CCGGCGGGCCCGCGCGCTCTTCCGCCGCTCTCGCATGCCCATGGCCGGCCGGTCC CACCCGGGCCCCGTGCGGGGCGGCGCGTGTGCTCTCTTGTGGTGGCCGCGTGCCT CCTGCCCGGAGCCGGCGGGACATGCCCGGAGCGCGCGTGGAGCGGCGGAGGAGGAG GCGAACGTGGTGCTACCGGGACGGTGGAGGAGATCCTCAACGTGGACCCGGTGCAGC ACACGTACTCCTGCAAGGTTCCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGTGGACGGCGGCAACAAGGTGGTGATCAGCGGCTTTGGAGACCCC CTCATCTGTGACAACAGGTGTCCACTGGGGACACCAGGATCTTCTTTGTGAACCCCTG CACCCCATACCTGTGGCCAGCCCAAGAACGAGCTGATGCTCAACTCCAGCCTCAT GCGGATCACCTGCGGAACCTGGAGGAGGTGGAGTTCTGTGTGGAAGATAAACCCGGG ACCCACTTCACTCCAGTGCTCCGACGCTCTGTATGCGTGCCGGGAATGCTGTGCG GCTTCGGCGCGGTGTGCGAGCCCAACGCGGAGGGGCGGGCGGGCGTCTGCGTCTG CAAGAAGAGCCCGTGCCCGAGCGTGGTGGCGCTGTGTGTGGGTGCGACGCTCCACC TACAGCAACGAATGCGAGCTGACGCGGGCGCAGTGACGACGAGCAGCGCCGATCCGCC TGCTCAGCCGCGGGCGGTGCGGCTCGCGGACCCCTGCTCCAACGTGACCTGACGCTT CGGCAGCACCTGTGCGCGCTCGGCGGACGGGCTGACGCGCTCGTGCCTGTGCCCGCG ACCTGCCGTGGCGCCCCGAGGGGACCGTCTGCGGACGCGAGCGCGCCGACTACCCCG GCGAGTGCCAGCTCTGCGCGCGCTGCGCCCGCAGGAGAATGTCTTCAAGAAGTT CGACGGCCCTTGTGACCCCTGTGAGGGCGCCCTCCCTGACCGGACGCTGCGCTG GTGAACCCGCGCACGCGCGCCCTGAGATGCTCTACGGCCGAGAGCTGCCCTGCC GGCAGCGCCAGTGTGTGGGGACGACGAGTCACTACGAAACGACTGTGTATGCGG CCGATCGGGGGCGCCCGGGGTCTCTCTGCGAGAAAGTGCGCTCGGCGCAGTGCCAG GGTGAGACCAAGTGCCCGGAGCCCTGCCGGTTCAATGCCGTGTGCTGTCCCGCGTG GCCGTCCCGCTGTCTCTGCGACCGCGTCACTGTGACGGGGCTTACAGGCCCGTGTG TGCCAGGACGGGCGCAGTATGACAGTGATTGCTGGCGGCGAGGCTGAGTGCCCGG CAGCAGCGTGCCATCCCGAGCAAGCACCAGGGCCCGTGTGACCGCCCGCTGCCAT GCCTCGGGGTGAGTGTGCTTTGGGGCGACGTGTGCTGTGAAGAAGCGGCGAGGAGC GTGTGAATGCTGACGGCGTGTGAGCCTCTACGATCCTGTGTGCGGACGACGAGC GTCACATACGCGAGCGCTGCGAGCTGGAGGCCACGGCTGTACCTCGGGCGGAGA TCCAGGTGGCGCGCAAAGGACCTGTGACCGCTGCGGGACGTCGCGCTTTGGAGCCCT GTGCGAGGCCGAGACCGGGCGCTGCGTGTGCCCTCTGAATGCGTGGCTTTGGCCAG CCCGTGTGTGCTCCGACGGGCGACGTAACCCAGCGAGTGATGCTGACGTTGCACG CCTGCACACACAGATCAGCTGACGCTGCGCTCAGCTGGACCTGTGAGACCTGTGG AGATGCCGTGTGTGCTTTTGGGGCTGTGTGCTCCGCGGGGAGTGTGTGTCTCCCG TGTGAGCACCCCGCCCGGCCCCGTGTGTGGCAGCGACGCTGTCACTACGGCAGTG CCTGCGAGCTACGGGAAGCCGCTGCCCTCCAGCAGACAGATCGAGGAGGCCGGGC AGGGCCGTGCGAGCAGGCCGAGTGCCTTCCGGAGGCTGTGGCTTGGGAGGACGGT GACTGTGAGCAGGAGCTGTGCCGGCAGCGCGTGGCATCTGGGACGAGGACTCGGAGG ACGGGCCGTGTGTCTGTGACTTCACTGTCCAGAGTGTCCAGGCGAGCCCGTGTGCGG CTAGATGGGGTCACTACAGCACCGAGTGTGAGCTGAAGAAGCCAGGCTGTGAGTCA CAGCGAGGGCTCTACGTAGCGGCCAGGGAGCCTGCCGAGGCCCGCTTCCGCCCGC TGCCGCTGTGGCCCCCTTACACTGTGCCAGACGCCCCACGGCTGTGCCAGGACAA TATCACCGCAGCCCGGGCGTGGGCTGGCTGGCTGCCAGTGCCTGCCAGTGCAAC CCCCATGGCTCTTACGGCGGCACTGTGACCCAGCCACAGGCCGCTGCTGCGGCC CAGGTGTGGGGGCTCAGGTGTGACCGCTGTGAGCCTGGCTTCTGGAATTTTCAGG CATCGTACCGATGGCGGAGTGGCTGTACACCTGACGCTGTGATCCCCAGGGCGCC GTGCGGATGACTGTGAGCAGATGACGGGGCTGTGCTGTGTAAGCCCGGGGTGGTG GACCAAGTGTGGGAGTGTCCAGACGGCGTGCCTGGGCCCCCGGGGCTGTGAAGC TGACGCTTCTGCGCTGCGACCTGTGCGGAGATGCGCTGTGAGTTCCGTGCGCGGTGC GTGGAGGAGTCTGGCTCAGCCCACTGTGTCTGCCCGATGCTCACTGTCCAGAGGCCA ACGCTACCAAGGTCTGTGGGTGAGATGGAGTCAATACGGCAACGAGTGTGAGCTGAA GACCATCGCCTGCCCGCAGGCGCTGCAATCTCTATCCAGAGCCTGGGCCCCGTGCCAG GAGGCTGTGTCTCCAGCACTACCCGACATCTGCTCCGTGACTGTGACCAACCCAG GGCTCTCTGAGCCAGGCACTGCCGGCCCCCCCCCGGCGCCCTCCCTGGCTCCAG CAGTACCGCACACAGCCAGACACCCCTCCGCCCTCATCGGACCTCGGACCTGCC AGCGTCCCAGGACACCGTGTGGCCCGTGTGACGGTCCCCCACGGCACCCCTCCC CTGACCCAGGCTGTGGCGTCCGCTTTGGTGAATCTGGCAGCACTGATGGAAGCAG CGATGAGGAAGTGAAGGGGACAGGAGGCCAGTGGGGTGGCTCTGGGGGTCTGAG CCCTTGGAGGGCAGCAGCTGGCCACCCCTGGGCCACCTGTGAGAGGGCTTCTGCT ACAACTCCGCGTGGGCTGTGCTCTGATGGGAAGACGCCCTCGCTGACGCGAGAGG GTCCAGGAGCTGTTCTACACGCCCGAGATGGCTGACCCCAAGTCAGAACTGTTCCGGG AGACAGCCAGGAGCATTGAGAGCACCTGGACGACCTCTCCGGAATTGACAGCTCAA GAAGGATTTCCGGAGTGTCCGCTTGGCGGACCTGGGGCCCCGCAATCCGTCCGCGCC ATTGTGGATGTGCACTTTGACCCACACAGCCTTCAGGGACCCGACGTGGCCCGGG CTTGCTCCGGCAGATCCAGGTGTCCAGGCGCGGTCTTGGGGGTGAGGCGCGCT GCAGGAGCAGTGCATTTATGGAATTTGACTGGTTTCTGCGCTTTATCACGGGGGCC ACGTCAGGAGCATTGTGCGGGAGCCAGGCGCAGAGCCCACTGCATCGCGCTGC </p>	

	<p>CGTCTCTGCTGTGACCCCTCGGGCCCCGACCCCACTCACAAAGCCAGCCCGTTGC CAAGACCACGGCAGCCCCACACAGTCGGCCCCCACCCTGCCCCAGCCGTGTG CCCCGACGTGGGCCCCGGCCCCCAGCAGCTCCAAAGCCCTGTGACTCACAGCCCT GCTTCCACGGGGGACCTGCCAGGACTGGGCATTGGCGGGGGCTTCACCTGCAGCTG CCCCGACGGCAGGGAGGCGCCGTCTGTGAGAAGGTGCTTGGCGCCCCCTGTGCCGGCC TTCGAGGGCGCTCCTTCTGGCCTTCCCCACCCTCCGCGCTTACCACAGCTGCGGCC TGGCACTGGAATTCGGGGCGCTGGAGCCTCAGGGGCTGTGTGTACAATGGCAACGC CCGGGGCAAGGACTTCTGGCATTTGGCGCTGTAGATGGCCGCTGCAGCTCAGGTTT GACACAGGTTCCGGGGCGGCGGTGTGACCACTGCGGTGCCGGTAGAGCCGGGCCAGT GGCACCGCTGGAGCTGTCCCGGCACTGGCGCCGGGGCACCTCTCGGTGGATGGTGA GACCCCTGTCTGGGCGAGAGTCCAGTGGCACCGAGCCCTCAACCTGGACACAGAC CTCTTTGTGGGGCGGCTACCCGAGGACCAGGCTGCCGTGGCGCTGGAGCGGACCTTCG TGGGCGCCGGCCTGAGGGGGTGCATCCGTTTGTGTGACGTCAACAAACAGCGCTTGA GCTTGGCATTGGGCGGGGGGTGCCACCCGAGGCTTGGCGTGGGCGAGTGGCGGGAC CACCCCTGCTGCCAACCCCTGCCATGGCGGGGGCCCATGCCAGAACCTGGAGGCTG GAAGGTTCCATTGCCAGTCCCGCCCCGGCGCGTGGACCAACCTGTGCCGATGAGAA GAGCCCTGCCAGCCCAACCCCTGCCATGGGGCGGCGCTGCCGTGTGTGCCCGAG GGTGGTGTCTAGTGCAGTGGCCCTTGGGGCGTGGGGCACCTTCTGCCAGACAGCCT CGGGGCGAGGCTTGGGGCCCTTCTGGCTGACTTCAACGGCTTCTCCACCTTGA GCTGAGAGGCTGCACACCATTCACGGGACCTGGGGGAGAAGATGGCGCTGGAGGCC GTGTTCTTGGCAGAGGCCCCAGCGGCTCCTGTCTACAACGGGCAAGAGCGGACG GCAAGGGGACTTCGTGTGCTGGCACTGGGGACCGCGCTGGAGTTCCGCTACGA CCTGGGCAAGGGGGCAGCGGTATCAGGAGCAGGGAGCCAGTCAACCTGGGAGCCTGG ACAGGGGTCTACTGGAGCGAAACGGCCGCAAGGGTGCCTGCGTGTGGGCGACGGCC CCCGTGTGTGGGGGAGTCCCCGAATCCCGCAAGGTTCCGCAACCGTCTCAACCT GAAGGAGCGCTCTACGTAGGGGGCGCTCCCGACTTCAAGCAAGCTGGCGTGTGT GCCGTGTCTCTGGCTTCGACGGTGCATCCAGCTGGTCTCCCTCGGAGGCGGCCAGC TGCTGACCCCGGAGCAGTGTGCGGCGAGTGGACGTCACTCTTTCAGAGTCACCC CTGCACCCGGGCTCAGGCCACCCCTGCCCTCAATGGGGCTCCTGCGTCCCGAGGGAG GCTGCCTATGTGTGCTGTGTCCCGGGGATTCTCAGGACCGCACTGCGAGAAGGGC TGGTGGAGAAGTCAGCGGGGACGTGGATACCTTGGCCTTTGACGGGCGGACCTTTGT CGAGTACCTCAACGCTGTGACCGAGAGCGAGAAGGCATGACAGCAACCACTTTGAA CTGAGCCTGCGCACTGAGGCCACGCAAGGGCTGGTGTCTGAGTGGCAAGGCCAGG AGCGGGCAGACTATGTGCACTGGCCATTGTGGACGGGCACCTGCAACTGAGCTACAA CCTGGGCTCCAGCCGCTGGTGTGCTGCGTCCACCGTGCCTCAACCAACCGCTGG TTGCGGGTCTGGCACATAGGGAGCAGAGGAAGGTTCCCTGCAGTGGGCAATGAGG CCCTGTGACCGGCTCCTCCCGCTGGGGCCACGAGCTGGACACTGCGAGAAGGCTG GTGGCTTGGGGGCTGCCGGAGCTGCCCGTGGGGCCAGCACTGCCAAGGCTACGGC ACAGGCTTTGTGGGCTGCTTGGCGGATGTGGTGGTGGGGCCGACCCGCTGCACCTGC TGGAGGACGCGCTACCAAGCCAGAGCTGCGGCCCTGCCCCACCCATGAGCTGGCAC CAGAGCCCCGCGCCGCT</p>
	<p>ORF Start: ATG at 37</p>
	<p>ORF Stop: TGA at 6196</p>
	<p>SEQ ID NO: 96</p>
	<p>2053 aa</p>
	<p>MW at 215628.0kD</p>
<p>NOV32a, CG94946-01 Protein Sequence</p>	<p>MRHGRPVPPGPAAGRPLPLLVVAACVLPAGAGTGPALERLREREEANVLTGTVEBI LNDVQVQHTYSCKVRVWRYLKGKDLVARESLDGGNKVVISGFDPLICDNQVSTGDT RIFVNPAPPYLPWPAHKNELMLNSSLMRI TLRLNEEVEFCVEDKPGTHFTVPPTPD ACRGMLCGFAGVCEPNAEGPGRASCVCCKSPCPVVAPOGSDASTYSNECELQRAQC SQQRIRLLSRGPCSRDPSCNVTCSTFSGTSCARSADGLTASCLCPATCRGAPEGTVC SGADYPGECQLLRACARQENVFKFDGPDPCQALPDPSPRSRNVPRTRRPEMLL RPESCPARQAPVCGDDGVTYENDCVMGRSGAARGLLQKVRSGQCQRDQCEPCRFRN AVCLSRGRPRSCDRVTCDGAYRPVCAQDGRYDSDCWRQAEQRQRAIPSKHQGP CDQAPSPCLGVQCAFATCAVKNGQAACECLQACSSLYDPVCGSDGVTYGSACELEAT ACTLGREIQVARKGPCDRGQCRFGALCEAETGRCVCPSECVALAQPVCSDGHTYPS ECMLHVHACTHQISLHVASAGPCETCGDAVCAFGAVCSAGQVCPCREHPPPGVCGS DGVTYGSACELREAACLTQQTQIEEARAGPCEQAECSGSGSGEDDCEQELCRQRGG IWDESEDGPCVDFSCQSVPGSPVCGSDGVTYSTECCLKKARCESQRLYVAAQAGC RGPAPFAPLPVAPLHCAQTPYGCCQDNITAARGVGLAGCPSACQCNPHGSYGGTCDPA TGQCSRCRPGVGLRCDRCPEGFWNFRGIVTDGRSGCTPCSCDPQGAVRDDCEQMTGLC SCKPGVAGPKCGQCPDGRALGPAGCEADASAPATCAEMRCEFGARCEVESGSAHVCPC MLTCEANATKVCSDGVTYGNECQLKTLACRQLQISIQSLGPCQEAFAVAPSTHTSA SVTVTTPLLLSQALPAPPALPLAPSSTAHSQTPPPSSRPRTTASVPRTTVWPVLT VPPTAPSPAPSLVAFGESGSDTGSSEELSGDQEAEGGGSGGLEPFEGLSSVATPGP PVERASCYNALGCCSDGKTPSLDAEGSNCPATKVFQGVLELEGVEGQELFYTPEMAD PKSELFGETARSIESTLDDLFRNSDVKKDFRSVRLRDLGPGKSVRAIVDVHFDPTTAF RAPDVARALLRQIQVSRRLSLGVRRPLQEHVRFMDFDWFPATITGATSGAIAAGATAR ATTASRLPSSAVTPRAPHPSHTSQFVAKTTAAPTTRRPTTAPSPVGRPPAPQPPP KPCDSQPCFHGGTCQDVALGGGFTCSCPAGRGGAECVLAGAPVPAFEGVCSFLAFTL RAYHTLRLALEFRALEPQGLLLYNGNARGKDFLALALLDGRVQLRFDTGSGPAVLTS VPVEPGQWHRLSRHWRRTLSVDGETPVLGESPSGTDGLNLDLDFVGGVBPEDQAA VALERTFVGAGLRGICIRLLDVNNQRLLELIGPGAATRGSGVGECDHPCLPNPCHGGA PCQNLEAGRFHCQCPGRVGPPTCADEKSPCQPNPCHGAAPCRVLPPEGGAQCECPLGRE</p>

	<p>GTFCQTASGQDGSPPFLADFNFSHLELRGLHTIARDLGEKMALEAVFLARGPSSGLLL YNGQKTDGKGFVSLALRDRRLEFRYDLGKGAIVRSREPVTLGAWTRVSLERNRKG ALRVGDGPRVLGESPKSRKVPHTVLNLKEPLYVGGAPDFSKLARAAVSSGFDGAIQL VSLGGRQLLTPEHVLRLQVDVTSFAGHPCTRASGHPCLNGASCVPREAYVCLCPGGFS GPHCEKGLVEKSAGDVTDLAFDGRFTVEYLNVAVTESEKALQSNHPELSLRTEATQGLV LWSGKATERADYVALAIVDGHLLQLSYNLGSQPVVLRSTVPVNTNRWLRVVAHREOREG SLQVGENAPVTGSSPLGATQLDLDGALWLGGLPELPGPALPKAYGTGFGVGLRDVVV GRHPLHLLLEDAVTKPELRPCPTP</p>	
	SEQ ID NO: 97	4760 bp
NOV32b, CG94946-02 DNA Sequence	<p>CCGGCGCGGCCCGCGCGCTCTTCCGCGCGCTCTCGCATGCGCCATGGCCGGCCGGTCC CACC CGGCGCCGCTGCGGGCGCGCGCTGCTGCCTCTCCTTGTTGGTGGCCGCTGCGT CCTGCCCGGAGCCGCGGGACATGCCCGGAGCGCGCGCTGGAGCGCGCGAGGAGGAG GCGAACGTGGTGTCTACCGGGACGCTGGAGGAGATCCTCAACGTGGACCGCGTGCAGC ACACGTACTCTGCAAGGTTCCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGGCGGCAACAAGGTGGTGTATCAGCGGCTTTGGAGACCCC CTCATCTGTGACAACAGGTGTCCACTGGGGACACCAGGATCTTCTTTGTGAACCTTG CACCCTCATACCTGTGGCCAGCCCAAGAAGAGCTGTATGCTCAACTCCAGCCTCAT GCGGATCACCCTGCGGAACCTGGAGGAGGTGGAGTTCTGTGTGGAAGATAAACCGGG ACCCACTTCACTCCAGTGCCTCCGACGCTCTGTATGCTGCGGGGAATGCTGTGCG GCTTCCGGCGCGGTGTGCGAGGCCAACGCGGAGGGGCCGGGCCGGCGCTCTGCGTCTG CAAGAAGAGCCCGTGCCTCCAGCGTGGTGGCGCTGTGTGTGGTGGAGCGCTCCACC TACAGCAACGAATGCGAGCTGCGAGCGGCGCAGTGCAGCCAGCAGCGCATCCGCGC TGCTCAGCGCGGGCGGTGCGGCTCGCGGGACCCCTGCTCCAACGTGACCTGCAGCTT CGGCAGCACCTGTGCGCGCTCGGCCGACGGGCTGACGGCTCGTGCCTGTGCCCCGCG ACCTGCCGTGGCGCCCCGAGGGACCGTGTGCGCGCAGCAGCGCGCGACTACCCCG GCGAGTGCCAGCTCTGCGCGCGCGCTGCGCCGCGCAGGAGAATGTCTTCAAGAAGTT CGAGGGCCCTTGTGACCCCTGTGAGGGCGCCCTCCCTGACCCGAGCGCGAGCTGCCGT GTGAACCGCGCAGCGCGCGCCCTGAGATGCTCTACGGCCGAGAGCTGCCCTGCC GGCAGGCGCCAGTGTGTGGGGACGACGGAGTCACTACGAAACGCACTGTGTATGGG CCGATCGGGGGCGCCCCGGGTCTCCTCTGCAAGAAAGTGGCTCCGGCCAGTGCCAG GGTGCAGACCACTGCCCGAGCCCTGCCGTTCAATGCCGTGTGCTGTGCCCGCGTGTG GCCGTCCCGCTGTCTCTGCGACCGCTCACCTGTGACGGGCGCTACAGGCCCGTGTG TGCCCGAGACGGGCGCAGTATGACAGTGTATGCTGGCGCGCAGCGCTGAGTGCCCG CAGCAGCGTGCCATCCCGAGCAAGCACCAGGGCCCGTGTGACAGGCCCGCTCCCAT GCTTCCGGGTGCGAGTGTGATTTGGGGCGAGTGTGTGTGAAGAACGGGCGAGCAGC GTGTGAATGCCCTGCGAGCGGTGTCTGAGCCCTACAGATCTGTGTGCGGCGAGCAGCG GTCACATACGGCAGCGGTGCGAGCTGGAGGCCAGGGCTGTACCCTCGGGCGGGAGA TCCAGGTGGCGCGCAAGGACCCCTGTGACCGCTGCGGGCAGTGCCTTTGGAGCCCT GTGCGAGGCGAGACCGGGCGCTGCGTGTGCCCCCTGTAATGCGTGGCTTTGGCCAG CCCGTGTGTGGTCCGAGCGGCGACAGTACCCAGCGAGTGCATGCTGCAGTGCACG CCTGCACACACCAGATCAGCCTGCAGTGGCTCAGCTGGACCTGTGAGACCTGTGG AGATGCCGTGTGTGCTTTTGGGGCTGTGTGCTCCGCGAGGCGAGTGTGTGTGTCCCG TGTGAGCACCCCGCGCCCGCGCCCGTGTGTGGCAGCGAGCGTGTACCTACGGCAGTG CCTGCGAGCTACGGGAAGCCGCTGCTCCAGCAGACAGATCGAGGAGGCCCGGGC AGGGCCGTGCGAGCAGGCGGAGTGGGTTCCGGAGGCTCTGGCTCTGGGGAGGACGCT GACTGTGAGCAGGAGCTGTGCGGCGCGCGTGGCATCTGGGACGAGGACTCGGAGG ACGGGCCGTGTGTGTGACTTACGTGCGCAGAGTGTCCAGCGAGCCCGGTGTGCGG CTCAGATGGGGTCACTACAGCAGGAGTGTGAGCTGAAGAAGGCCAGGTGTGAGTCA CAGCGAGGGCTTACGTAGCGGCCAGGAGGCTGCGGAGGCCCGGCTTCCGCCCCG TGCCGCTGTGGCCCCCTTACACTGTGCCAGACGCCCTACGGCTGTGCCAGGACAA TATCACCAGCAGCCGGGGCGTGGGCTGGTGGCTGCCCCAGTGCCCTGCCAGTGCAAC CCCCATGGCTCTTACGGCGGCACCTGTGACCCAGCCACAGGCCAGTGTCTCTGCCGCC CAGGTGTGGGGGGCTCAGGTGTGACCGCTGTGAGCCTGGCTTCTGGAACCTTCGAGG CATCGTCAACGATGGCGGAGTGGCTGTACACCTGACGTGTGATCCCAAGGCGCC GTGCGGGATGACTGTGAGCAGATGACGGGGCTGTGCTCGTGTAAAGCCCGGGTGGCTG GACCCAGTGTGGGCGAGTGTCCAGACGGCGGTGCCCTGGGCCCGCGGCTGTGAAGC TGACGCTTCTGCGCTGCGACCTGTGCGGAGATGCGCTGTGAGTTCGGTGGCGCGTGC GTGGAGGAGTGTGGCTCAGCCCACTGTGTGTGCTGCGGATGCTCACTGTCCAGAGGCCA ACGCTACCAAGTCTGTGGGTGAGTGGAGTACATACGGCAACGAGTGTGAGTGA GACCATCGCTGCGGCCAGGGCCTGCAATCTCTATCCAGAGCTGGGCCGTGCCAG GAGGCTGTGTGCTCCAGCACTCACCCGACATCTGCTCCGTGACTGTGACCAACCCAG GGCTCCTCTGAGCCAGGCACTGCGGGCCCCCCCCCGGGCGCCCTCCCCCTGCCAG CAGTACCGCACACAGCCAGACCCCTCCGCCCTCATCGGACCTCGGACCACTGCC AGCGTCCCCAGGACACCGTGTGGCCCGTGTGACCGTGGCCCCCAGGCAACCTCCC CTGCACCTCAGCTGTGGCGTCCGCTTTGGTGAATCTGGCAGCATGTGAAGAGCAG CGATGAGGAATGAGCGGGGACAGGAGGCCAGTGGGGTGGCTCTGGGGGCTCGAG CCCTTGAGGGGCGAGCGTGGCCACCCCTGGGCCACCTGTGAGAGGGCTTCTGTCT ACAACCCCTGCCATGGGGCGGCGCCCTGCGGTGTGCTGCCGAGGGTGGTGTCTAGTG CGAGTGCCCCCTGGGGCGTGGGGCACCTTCTGCCAGACAGCTCGGGGCGAGACGGC TCTGGGCCCCCTCTGTGCTGACTTCAACGGCTTCTCCACCTGGAGCTGAGAGGCTGC ACACCATGACCGGACCTGGGGGAGAAGATGGCGCTGGAGGCCGTGTCTGTGGCAGC AGGCCCGAGCGGCTCTGTCTACAAACGGGCGAAGACGGGCGGCAAGGGGACTTC</p>	

	GTGTGCTGGCACTGCGGGACCGCCGCTGGAGTTCCGCTACGACCTGGGCAAGGGG CAGCGGTTCATCAGGAGCAGGGAGCCAGTCACCTGGGAGCCTGGACCAGGGTCTCACT GGAGCGAAACGGCCGCAAGGGTGCCCTGCGTGTGGGCGACGGCCCCGCTGTGTGGGG GAGTCCCCGAAATCCCGCAAGGTTCCGCACACCGTCTCAACCTGAAGGAGCCGCTCT ACGTAGGGGGCGCTCCCGACTTCAGCAAGCTAGGCGCGTGTCTGCTGCGTCTCTGG CTTCGACGGTGCCATCCAGCTGGTCTCCCTCGGAGGCCGCGCAGCTGCTGACCCGGAG CACGTGTCTGGCGCAGGTGGACGTACAGTCTCTTTGACAGGTACCCCTGCACCCGGGCTC CAGGCCACCCCTGCCCTCAATGGGGCCTCTGCGTCCGAGGGAGGCTGCCATGTGTG CCTGTGTCCCGGGGATTCTCAGGACCGCACTGCGAGAAGGGGCTGGTGGAGAAGTCA GCGGGGGACGTGGATACCTTGGCCTTTGACGGGCGGACCTTTGTGAGTACCTCAAGC CTGTGACCGAGCGGAGAAGGCACTGCAGAGCAACCACTTTGAAGTGAAGCTGCGCAC TGAGGCCACGCAAGGGGCTGGTGTCTTGGAGTGGCAAGGCCAGGAGCGGGCAGACTAT GTGGCACTGGCCATTGTGGACGGGCACCTGCAACTGAGCTACAACCTGGGCTCCCAGC CCGTGGTGTGCGTTCCACCGTGCCCGTCAACACCAACCGCTGGTGTGGGTCGTGGG ACATAGGGAGCAGGGGAAGGTTCCCTGCAAGTGGGCAATGAGGCCCTGTGACCGGC TCCTCCCCGCTGGGCGCCACGCACTGGACACTGATGGAGCCTGTGGCTTGGGGGCC TGCCGGAGCTGCCCCTGGGCGCCAGCACTGCCCAAGGCCTACGGCACAGGCTTTGTGGG CTGCTTGGGGATGTGGTGGGCGGCAACCGCTGCACCTGTCTGGAGGACGCGCTC ACCAAGCCAGAGCTGCGGCCCTGCCCAACCCATGAGCTGGCACAGAGCCCGCGCC CGCT		
	ORF Start: ATG at 37	ORF Stop: TGA at 4732	
	SEQ ID NO: 98	1565 aa	MW at 163817.1kD
NOV32b, CG94946-02 Protein Sequence	MRHGRPVPPGPAAGRPLLPLLVAAACVLPAGAGTCTPERALERREEANVLTGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLLDGGNKVVISGFGDPLICDNQVSTGDT RIFFFNPAPPYLVPAHKNELMLNSSLMRITLRNLEEVEFCVEDKPGTHTPTVPPTPPD ACRGMLCGFAGVCEPNAEGPGRASCVCCKSPCPSVVPVCGSDASTYSNECELQRAQC SQRRIRLLSRGPCGSRDPCSNVTCSTFGSTCARSADGLTASCLCPATCRGAPGTGVC SDGADYPGECQLLRRACARQENVFKFDGPCDPCQALPDPSPSRSCRVPNPRTRPEMLL RPESCPARQAPVCGDDGVTYENDCVMGRSGAARGLLLQKVRSGQCGRQDQCPPEPCRFN AVCLSRGRPRCSDRVTCGAYRPVCAQDGRYDSDCWQQAECRQRAIPSKHQGP CDQAPSPCLGVQCAFAGTCAVKNQQAACECLQACSSLYDPVCGSDGVTYGSACELEAT ACTLGREIQVARKGPCDRCGQCRFGALCEAETGRVCVPCSECVLAQPVCGSDGHTYPS ECMLHVHACTHQISLHVASAGPCETCGDAVCAFGAVCSAGQCVPCREHPPPGVCGS DGVTYGSACELEAACECLQQTQIEEARAGPCEQAECGSGGSGSGEDGDCQELCRQGG IWDESEDGPGCVCFDSCQSVPGSPVCGSDGVTYSTECLEKARCESQRLVYAAQGAC RGPAPFAPLPVAPLHCAQTPYGCCQDNITARGVGLAGCPSACQCNPHGSYGGTCDPA TGQCSRCRPGVGLRLCDRCEPGFWNFRGIVTDGRSGCTPCSCDPQGAVRDDCEQMTGLC SCKPGVAGPKCGQCPDGRALGPAGCEADASAPATCAEMRCEFGARCVESGSAHCVCP MLTCEANATKYCGSDGVTYGNELQKTIACRQGLQISIQSLGPCQEAVALPSTHTPTSA SVTVTTPGLLSQALPAPPALPLAPSSTAHSQTTTPPSSRPRTTASVPRTTVWPVLT VPPTAPSPAPSLVASAFGESGSTDGGSSDEELSGDQEAAGGGSGGLEPLEGGSSVATPGP PVERASCYNPCHGAAPCRVLPAGGAQCECPLGREGTFCQTASGDGSGPFLADFNFGS HLELRGLHTIARDLGEKMALEAVFLARGPSGLLLYNGQKTDGKGFVLSIALDRRLRF RYDLGKGAIVIRSREPVTLAGWTRVSLERNRKGALRVGDGPRVLGESPKSRKVPHTV LNLKEPLYVGGAPDFSKLAAAAVSSGFDGAIQLVSLGGRQLLTPHVLRLQVDVTSFA GHPCTRASGHPCLNGASCVPREAAVYVCLCPGGFSGPHCEKGLVEKSGADVDTLAFDGR TFVEYLNATVESEKALQSNHFLSLRTEATQGLVLWSGKATERADYVALAIVDGHLLQ SYNLGSQPVVLRSTVPVNTNRWLRVVAHREQREGSLQVGNAPVTVSSPLGATQLDLD GALWLGLPELPGPALPKAYGTGFGVCLRDVVVGRHPLHLEDAVTKPELRPCPTP		
	SEQ ID NO: 99	893 bp	
NOV32c, CG94946-03 DNA Sequence	CCGGCGCGGCCCGCGCGCTCTTCGCGCGCTCTCGCATGCGCCATGGCCGGCCGGTCC CACCCGGGCCCGCTGCGGGGCGGCCGCTGCTGCCTCTCCTTGTGGTGGCCGCGTGCCT CCTGCCCGGAGCCGGCGGGACATGCCGAGCGCGCTGGAGCGCGCGGAGGAGGAG GCGAACGTGGTGTCTCACCGGGACGGTGGAGGAGATCCTCAACGTGAGACCGGTGCAGC ACACGTACTCTGCAAGGTTCCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGGCGGCAACAAGGTGGTGATCAGCGGCTTTGGAGACCCC CTCATCTGTGACAAACAGGTGTCCACTGGGGACACAGGATCTTCTTGTGAACCTG CACCCCATACCTGTGGCCAGCCACAAGAACGAGCTGGTGTCTTGGAGTGGCAAGGC CACGGAGCGGGCAGACTATGTGGCACTGGCCATTGTGGACGGGCACTGCAACTGAGC TACAACCTGGGCTCCAGCCCGTGGTGTGCTGCGTTCCACCGTGCCCGTCAACACCAACC GCTGGTTGCGGGTCTGGGCACATAGGGAGCAGAGGGAAGGTTCCCTGCAAGGTGGGCAA TGAGGCCCTGTGACCGGCTCTCTCCCGCTGGGCGCCACGAGCTGGACACTGATGGA GCCCTGTGGCTTGGGGGCTGCCGGAGCTGCCCGTGGGCGCCAGCACTGCCAAGGCTC ACGGCACAGGCTTTGTGGGCTGCTTGGGGATGTGGTGGTGGGCGGCGCCGCTGCA CCTGTCTGAGGACGCGCTACCAAGCCAGAGCTGCGGCCCTGCCCAACCCATGAGCT GGCACAGAGCCCGCGCGCCGCT		
	ORF Start: ATG at 37	ORF Stop: TGA at 865	
	SEQ ID NO: 100	276 aa	MW at 29688.8kD

NOV32c, CG94946-03 Protein Sequence	MRHGRPVPPGPAAGRPLLP LLVVAACVLPAGAGGTC PERALERREEEANVVL TGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLDGGNKVVISGFGDPLICDNQVSTGDT RIFFVNPAPPYLWPAHKNELVLWSGKATERADYVALAIVDGHQLQSYNLGSQPVVLR TVPVNTNRWL RVVAHREQREGSLQVGNAPVTGSSPLGATQLD TDGALWGLGPELPV GPALPKAYGTGFVGC LRDVVVGRHPLHLL EDAVTKEPRLPCPTP
--	--

	SEQ ID NO: 101	1931 bp
NOV32d, CG94946-04 DNA Sequence	CCGGCGCGGCCCGCGCGCTCTTCCGCCGCTCTCGCATGCGCCATGGCCGGCCGGTCC CACCCGGGCCCGCTGCGGGGCGGCCGCTGCTGCCTCTCCTTGTGGTGGCCCGCTGCGT CCTGCCCGGAGCCGGCGGGACATGCCCGGAGCGCGCTGGAGCGCGCGAGGAGGAG GCGAACGTGGTGCTCACCGGGACGGTGGAGGAGATCCTCAACGTGGACCCGGTGCAGC ACACGTACTCTGCAAGGTTCCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGGCGGCAACAAGGTGGTGATCAGCGGCTTGGAGACCCC CTCATCTGTGACAACAGGTGTCCACTGGGGACACCAGGATCTTCTTGTGAACCTG CACCCCATACCTGTGGCCAGCCCAAGAACAGAGCTGATGCTCAACTCCAGCCTCAT GCGGATCACCTGCGGAACCTGGAGGAGGTGGAGTTCTGTGTGAAGATAAACCCGGG ACCCACTTCACTCCAGTGCCTCCGACGCTCTGATGCGTCCGGGGAATGCTGTGCG GCTTCGGCGCCGTGTGCGAGCCCAACGCGGAGGGGCGGGCCGGCGCTCTGCGTCTG CAAGAAGAGCCCGTGCCCGAGCGTGGTGGCGCTGTGTGGGTGGAGCGCTCCACCC TACAGCAACGAATGCGAGCTGCGCGGGCGCAGTGCAGCCAGCAGCCGATCCGCC TGCTCAGCCGCGGGCGGTGCGGCTCGCGGGACCCCTGCTCAACGTGACCTGAGCTT CGGCAGCACCTGTGCGCGCTCGGCCGACGGGCTGACGGCTCTGTGCTGTGCCCGCG ACCTGCCGTGGCGCCCCGAGGGGACCGTCTGCGGCAGCGAGCGCGCGACTACCCG GCGAGTGCAGCTCCTGCGCCGCGCTGCGCCCGCAGGAGAAATGTCTCAAGAAAT CGACGGCCCTTGTGACCCCTGTGAGGGCGCCCTCCCTGACCCGAGCCGAGCTGCCGT GTGAACCCGCGCAGCGCGCCCTGAGATGCTCTACGGCCCGAGAGCTGCCCTGCC GGCAGGCGCCATGTGTGGGGACGACGGAGTCACTACGAAAACGACTGTGTATGGG CCGATCGGGGGCGCCCGGGGTCTCTCTGCGAGAAAGTGCGCTCCGGCCAGTGCCAG GGTGAGACCACTGCGCCGAGCCCTGCCGTTCAATGCCGTGTGCTGTCCCGCGTG GCGTCCCGCTGCTCTGCGACCGCGTCACTGTGACGGGGCTACAGGCCGCTGTG TGCCAGGACGGGCGCAGTATGACAGTGATTGCTGGCGGCGAGGCTGAGTGGCGG CAGCAGCGTGCCATCCCGAGCAAGCAGGAGGCGCGTGTGACAGGCCCGTCCCAT GCCTCGGGGTGAGTGTGATTTGGGGCGAGTGTGCTGTGAAGAAGGGCAGGACG GTGTGAATGCCTGACGGCTGCTCGAGCCTCTACGATCCTGTGTGCGGCGAGCAGG GTACATACGGCAGCGCTGCGAGCTGGAGGCCAGCGCTGTACCTCGGGCGGGAGA TCCAGGTGGCGCGCAAGGACCCCTGTGACCGCTGCGGGCAGTCCGCTTTGGAGCCCT GCCTGTGACCGGCTCTCCCGCTGGGCGCCACGAGCTGGACACTGAGGAGCCCTG TGGCTTGGGGGCTGCGGAGCTGCCCGTGGGGCCAGCACTGCCAAGGCCCTACGGCA CAGGCTTTGTGGGCTGCTTGCGGGATGTGGTGGTGGGCGGGCAGCCGCTGCACCTGCT GGAGGACGCCGTACCAAGCCAGAGCTGCGGCCCTGCCACCCCATGAGCTGGCACC AGAGCCCCGCGCCGCT	
	ORF Start: ATG at 37	ORF Stop: TGA at 1903
	SEQ ID NO: 102	622 aa MW at 66353.9kD
NOV32d, CG94946-04 Protein Sequence	MRHGRPVPPGPAAGRPLLP LLVVAACVLPAGAGGTC PERALERREEEANVVL TGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLDGGNKVVISGFGDPLICDNQVSTGDT RIFFVNPAPPYLWPAHKNELMLNSSLMRITLRNLEEVEFCVEDKPGTHFTVPPTPPD ACRGMLCGFGAVCEPNAEGPGRASCVCCKSPCPVVPVCGSDASTYSNECELRQAQC SQQRRI RLLSRGPCSRDPCSNVTCSFGSTCARSDGLTASCLCPATCRGAPEGTVCG SDGADYPGECQLLRACARQENVFKKFDGPCDPCQALPDPSSRSCRVPNPRTRRPEMLL RPESC PARQAPVCGDDGVTYENDCVMGRSGAARGLL LQKVRSGQCQGRDQCPEPCRFN AVCLSRGRPRCSDRVTCDGAYRPVCAQDGRTYDSDCWRQQAECRQRAIPSKHQGP CDQAPSPCLGVQCAFATCAVKNGQAACECLQACSSLYDPVCGSDGVTYGSACELEAT ACTLGREIQVARKGPCDRGQCRFGALPVTGSSPLGATQLD TDGALWGLGPELPVGP ALPKAYGTGFVGC LRDVVVGRHPLHLL EDAVTKEPRLPCPTP	

	SEQ ID NO: 103	4697 bp
NOV32e, CG94946-05 DNA Sequence	<p>CCGGCGGGCCCGCGCTCTTCCGCGCCTCTCGCATGCGCCATGGCCGGCCGGTCC CACCCGGGCCCCGCTGCGGGGCGGCGCTGCTGCCCTCTCCTTGTGGTGCCGCGTGGCT CCTGCCCGGAGCCGGCGGGACATGCCCGGAGCGCGCTGGAGCGGCGGAGGAGGAG GCGAACGTGGTGTCTACCGGGACGGTGGAGGAGATCCTCAACGTGGACCCCGTGCAGC ACACGTACTCCTGCAAGGTTGCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGGCGGCAACAAGGTGGTGATCAGCGGCTTTGGAGACCCC CTCATCTGTGACAACAGGTGTCCACTGGGGACACCAGGATCTTCTTTGTGAACCTTG CACCCCATACCTGTGGCCAGCCCAAGAACGAGCTGATGCTCAACTCCAGCCTCAT GCGGATCACCTGCGGAACCTGGAGGAGTGGAGTTCTGTGTGAAGATAAACCCGGG ACCCACTTCACTCCAGTGCCTCCGACGCCCTCTGATGCGTGGCGGGGAATGCTGTGCG GCTTCGGCGCCGTGTGCGAGCCCAACGCGGAGGGGCGGGCCGGGCGCTCTGCGTCTG CAAGAAGAGCCCGTCCCCAGCGTGGTGGCGCCTGTGTGTGGGTGGAGCGCTCCACC TACAGCAACGAATGCGAGCTGCAGCGGGCGAGTGCAGCCAGCAGCCCGCATCCGCC TGCTCAGCCGCGGGCGTGGCGCTCGCGGACCCCTGCTCCACGTGACCTGCAGCTT CGGCAGCACCTGTGCGCGCTCGGCCGACGGGCTGACGGCTCGTGCTGTGCCCCGCG ACCTGCCGTGGCGCCCCGAGGGGACCGTCTGCGGCAGCGAGCGGCCGACTACCCCG CGAGTGTCCAGCTCTGCGCGCGCCTGCGCCCCCAGGAGATGTCTTCAGAAAGTT CGACGGCCCTTGTGACCCCTGTGAGGGCGCCCTCCCTGACCCGAGCCGAGCTGCCGT GTGAACCCGCGCAGCGCGGCCCTGAGATGCTCTACGGCCCGAGAGCTGCCCTGCCC GGCAGGCGCCAGTGTGTGGGACGACGGAGTCACTACGAAACGACTGTGTGTCATGGG CCGATCGGGGCGCCCGGGGTCTCTCTGCGAGAAAGTGCCTCCGGCCAGTGCCAG GGTCGAGACCACTGCCCGGAGCCCTGCCGGTTCAATGCCGTGTGCTGTCCCGCGTG GCCGTCCCCGCTGCTCTGCGACCGCGTCACTGTGACGGGCGCTACAGGCCCGGTG TGCCAGGACGGGCGCAGTATGACAGTGTGCTGCGGCGCAGGCTGAGTGCCTCGG CAGCAGCGTGCCATCCCCAGCAAGCACCAGGGCCCGTGTGACCAGGCCCGCTCCCAT GCCTCGGGGTGCACTGTGATTTGGGGCGACGTGTGCTGTGAAGAACGGGCGAGGAGC GTGTGAATGCTGACGGCGTGTGAGCCCTACAGATCCTGTGTGCGGCAGCGACGGC GTCACATACGGCAGCGGTGCGAGCTGGAGGCCACGGCCTGTACCTCGGGCGGGAGA TCCAGGTGGCGCGCAAAGGACCCCTGTGACCGCTGCGGGCAGTGCCTTTGGAGCCCT GTGCGAGGCCGAGACCGGGCGCTGCGTGTGCCCTCTGAATGCGTGGCTTTGGCCCG CCCGTGTGTGGCTCCGACGGGCAACGTAACCCAGCGAGTGCACTGCGACGCTGCACG CCTGCACACACCAGATCAGCCTGCACGTGGCCTCAGCTGGACCTGTGAGACCTGTGG AGATGCCGTGTGTGCTTTTGGGGTGTGTGCTCCGCGAGGCGAGTGTGTGTCCCGG TGTGAGCACCCCGCCCGGCCCGTGTGTGGCAGCGACGGTGTCACTACGGCAGTG CCTGCGAGCTACGGGAAGCCGCTGCTCCAGCAGACACAGATCGACAGGCGCCGGC AGGGCCGTGCGAGCAGGCCGAGTGCCTTCCGGAGGCTCTGGCTCTGGGGAGGACGGT GACTGTGAGCAGGAGCTGTGCGGCGAGCGCGTGGCATCTGGGACGAGGACTCGGAGG ACGGGCGCTGTGTGACTTCACTGCGCAGAGTGTCCAGGACCGCGGTGTGCGGCT CTCAGATGGGGTCACTACAGCACCGAGTGTGAGCTGAAGAAGGCCAGGTGTGAGTCA CAGCGAGGGCTCTACGTAGCGGCCAGGGAGCCTGCCGAGGCCCGGCTTCCGCCCGC TGCGCCTGTGGCCCCCTTACACTGTGCCAGACGCCCTACGGCTGTGTCAGGACAA TATCACCGCAGCCCGGGCGTGGGCTGGCTGGCTGCCCGCAGTGCCTGCCAGTGCAAC CCCCATGGCTCTTACGGCGGCACCTGTGACCCAGCCACAGGCCAGTGTCTCTGCCGCC CAGGTGTGGGGGCGCTCAGGTGTGACCGCTGTGAGCCTGGCTTCTGGAATTTGAGG CATCGTCACCGATGGCCGAGTGGCTGTACACCTGACGTGTGATCCCCAAGCGGCC GTGCGGGATGACTGTGAGCAGATGACGGGCTGTGCTCGTGTAAAGCCGGGTGGCTG GACCCAAGTGTGGGCGAGTGTCCAGACGGCGGTGCCCTGGGCCCGCGGCTGTGAAGC TGACGCTTCTGCGCTGCGACCTGTGCGGAGATGCGCTGTGAGTTCCGTGCGCGGTGC TGGAGGAGTCTGGCTCAGCCCACTGTGTCTGCCGATGCTCACTCTCCAGAGGCCA ACGCTACCAAGTCTGTGGGTGAGTGGAGTACATACGGCAACGAGTGTGAGTGA GACCATCGCTGCCCGCAGGCCCTGCAATCTCTATCCAGAGCCTGGGCCCGTGCAG GAGGCTGTTGCTCCAGCACTACCCGACATCTGCTCCGTGACTGTGACCAACCCAG GGCTCTCTGAGCCAGGCACTGCCGGCCCCCGGCCCGCTCCCTGGCTGCCAG CAGTACCGCACACAGCCAGACCCCTCCGCCCTCATCGGACCTCGGACCACTGCC AGCGTCCCAGACACCCGTGTGGCCCGTGTGACGGTCCCCCAGGCAACCTCCC CTGCACCCAGCCTGGTGGCGTCCGCTTTGGTGAATCTGGCAGCAGTGAAGAGCAG CGATGAGGAATGAGCGGGGACAGGAGGCCAGTGGGGGTGGCTCTGGGGGCGCTGAG CCCTTGGAGGGGAGCAGCGTGGCCACCCCTGGGCCACCTGTGAGAGGGCTTCTCTGCT ACAACCTCCGCTTGGGTGCTGTCTGTGATGGGAAGACGCCCTCGTGGACGCGAGGG CTCCAACGCCCCCGCACCAAGGTGTTCAGGGCGTCTGAGCTGGAGGGCGTCCGAG GGCCAGGAGCTGTTCTACACGCCGAGATGGCTGACCCCAAGTCAAGACTGTTCCGGG AGACAGCCAGGAGCATGAGAGCACCTTGACGACCTCTTCCGGAATTCAGACGTCAA GAAGGATTTCCGGAGTGTCCGCTTGGGGACCTGGGGCCCGGCAATCCGTCCGCGCC ATTGTGATGTCACTTTGACCCCAACACAGCCTTCAGGGCAGCCGAGCTGGCGGGG CCCTGCTCCGGCAGATCCAGGTGTCCAGGCGCCGGTCTTGGGGGTGAGGCGGCCGCT GCAGGAGCAGTGGATTTATGGACTTTGACTGTTTCTGCGCTTTATCAGGGGGCC ACGTGAGGACATTGTGCGGGAGCCACGGCCAGAGCCACCACTGCATCGCGCTGC CGTCTGTGTGACCCCTCGGGGCCCGACCCCACTGCACACAAGCCAGCCCGTGTG CAAGACCAAGGAGCCCCACCAACGTGCGCCCCCACCCTGCCCCAGCCGTGTG CCCCGAGCTCGGCCCGGCTCTGCGTCCCGAGGGAGGCTGCTATGTGTGCTGT</p>	

	GTCCCGGGGGATTCTCAGGACCGCACTGCGAGAAGGGGCTGGTGGAGAAGTCAGCGGG GGACGTGGGATACCTTGGCCCTTTGACGGGCGGACCTTTGTGCGAGTACCTCAACGCTGTG ACCGAGAGCGAGAAGGCACTGCAGAGCAACCACTTTGAAGTACGCGCTGCGCACTGAGG CCACGCAGGGGCTGGTGCTCTGGAGTGGCAAGGCCACGGAGCGGGCAGACTATGTGGC ACTGGCCATTGTGGACGGGCACCTGCAACTGAGCTACAACCTGGGCTCCAGCCCGTG GTGCTGCGTTCCACCGTGCCCGTCAACACCAACCGCTGGTTGCGGGTCTGTGGCACATA GGGAGCAGAGGGAAGGTTCCCTGCAGGTGGGCAATGAGGCCCTGTGACCGGCTCCTC CCCGCTGGGCGCCACGCAGCTGGACACTGATGGAGCCCTGTGGCTTTGGGGGCTGCCG GAGCTGCCCCGTGGGCGCCAGCACTGCCCAAGGCCCTACGGCACAGGCTTTGTGGGCTGCT TGCGGGATGTGGTGGTGGGCCGGCACCCGCTGCACCTGCTGGAGGACGCCGTACACAA GCCAGAGCTGCGGCCCTGCCCCACCCCATGAGCTGGCACCAGAGCCCCGCGCCCCGT		
	ORF Start: ATG at 37		ORF Stop: TGA at 4669
	SEQ ID NO: 104		1544 aa MW at 162003.7kD
NOV32e, CG94946-05 Protein Sequence	MRHGRPVPPGPAAGRPLPLLVAAACVLPAGAGTGPERALERREEREEANVLTGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLLDGGNKVVISGFGDPLICDNQVSTGDT RIFFVNPAPPYLPWPAHKNELMLNSSLMRI TLRNLEEVEFCVEDKPGTHFTVPVPTPPD ACRGM LCGFAGVCEPNAEGPGRASCVCCKSPCSVVAPVCGSDASTYSNECELQRAQC SQQRRIRLLSRGPCGSRDPCSNVTCFSGSTCARSDGLTASCLCPATCRGAPEGTVCG SDGADYPGECQLLRACARQENVFKKFDGPCDPCQALPDPSSRCRVNPRTRRPEMLL RPESCPARQAPVCGDDGVTYENDCVMGRSGAARGLLLQKVRSGQCQRDQCFEPCRFN AVCLSRGRPRCSDRVTCDGAYRPVCAQDGRTYDSDCWRQAEACRQRAIPSKHQGP CDQAPSPCLGVQCAFATCAVKNGQAACECLQACSSLYDFVCGSDGVTYGSACELEAT ACTLGREIQVARKGPCDRCGQCRFGALCEAETGRCVCPSECVLAQPVCGSDGHTYPS ECMLHVHACTHQISLHVASAGPCETCGDAVCAFGAVCSAGQCVCPREHPPPGVPGS DGVTYGSACELEAACLQQTQIEEARAGPCEQAECGSGSGSGEDGDCQEQLCRQGG IWNDESEDGPCVDFSCQSVPGSPVCGSDGVTYSTECCLKKARCESQRLYVAAQAC RGPAFAPLPPVAPLHCAQTPYGCCQDNITAAAGVGLAGCPSACQCNPHGSYGGTCDPA TGQCSRCPGVGGLRCDRCPEPFWNFRGIVTDGRSGCTPCSDCPQAVRDDCEBQMTGLC SCKPGVAGPKCGQCPDGRALGPAGCEADASAPATCAEMRCEFGARCVEESGSAHCVCP MLTCEANATKVCSDGVTYGNECQLKTIACRQGLQISIQSLGPPQCAVAPSTHTPSA SVTVTTPGLLLSQALPAPPALPLAPSSSTAHSQTTPPPSSRPTTASVPRTTVWPVLT VPPTAPSPAPSLVASAFGESGSTDGSSEELSGDQEAASGGSGGLEPLEGSSVATPGP PVERASCYNALGCCSDGKTPSLDAEGSNCPATKVFQGVLELEGVGQELFYTPEMAD PKSELFGETARSIESTLDDLFRNSDVKKDFRSVRLRDLGPGKSVRAIVDVHFDPTTAF RAPDVARALLRQIQVRRRSLGVRRRLQEHVRFMDFFWPAFITGATSGAIAAGATAR ATTASRLPSSAVTPRAPHPSHTSQPVAKTTAAPTTRRPPTAPSRVPGRPPASCVP EAAYVCLCPGGFSGPHCEKGLVEKSAGDVTDLAFDGRFTFVEYLVNAVTESEKALQSNHF ELSLRTEATQGLVLWSGKATERADYVALAIVDGLQLSYNLGSPQVPLRSTVPVNTNR WLRVVHREQREGSLQVGNEAPVTGSSPLGATQLDLDGALWLGGLPELPVGPALPKAY GTGFVGLRDLVVVGRHPLHLEDAVTKPELRPCPTP		
	SEQ ID NO: 105		6494 bp
NOV32f, CG94946-06 DNA Sequence	CCGGCGCGGGCCCCGCGCTCTTCCGCGCCTCTCGCATGCGCCATGCGCGGCGCGGTCC CACCGGGGCCCGCTGCGGGGCGGCGCTGCTGCCTCTCCTTGTGGTGGCCGCGTGCCT CCTGCCCGGAGCGCGCGGGACATGCCCGGAGCGCGCTGGAGCGGCGCGGAGGAGG GCGAACGTGGTGTGCTCACCGGACGGTGGAGGAGATCCTCAACGTGGACCCCGTGCA ACACGTACTCTGCAAGGTTGCGGTCTGGCGGTACTTGAAGGGCAAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGGCGGCAACAAGGTGGTGTATCAGCGGCTTTGGAGACCCC CTCATCTGTGACAACCAAGGTGTCCACTGGGGACACCAGGATCTTCTTTGTGAACCTG CACCCCATACCTGTGGCCAGCCACAAGAACGAGCTGATGCTCAACTCAGCTCAGCTCAT GCGGATCACCTGCGGAACCTGGAGGAGGTGGAGTTCTGTGTGGAAGATAAACCCGGG ACCCACTTCACTCCAGTGCCCTCCGACGCCCTCTGATGCGTGCCGGGGAATGCTGTGG GCTTCCGGCGCGTGTGCGAGCCCAACGCGGAGGGGCGGGCCGGGCTGCTGCGTCTG CAAGAAGAGCCCGTGCCCGAGCGTGGTGGCGCCTGTGTGTGGGTGCGACGCCCTCCACC TACAGCAACGAATGCGAGCTGCAGCGGGCGCAGTGACGCCAGCAGCGCGCATCCGCC TGCTCAGCGCGCGGCGGTGCGGCTCGCGGACCCCTGCTCCAACGTGACCTGCAGCT CGGCAGCACCTGTGCGCGCTCGGCGGACGGGCTGACGGCTCTGCTGCTGCGCCCGG ACCTGCCGTGGCGCCCCGAGGGGACCGTCTGCGGCGAGCGAGCGCGGACTACCCCG GCGAGTGCCAGCTCCTGCGCGCGCGCTGCGCCCGCAGGAGAATGTCTTCAAGAAGTT CGACGGCCCTTGTGACCCCTGTGAGGCGCCCTCCCTGACCCGAGCCGCGAGCTGCCGT GTGAACCCGCGCACGCGGCGCCCTGAGATGCTCTACGGCCGAGAGCTGCCCTGCC GGCAGGCGCCAGTGTGTGGGGACGACGAGTCACTACGAAAACGACTGTGTATGGG CCGATCGGGGGCGCCCGGGGTCTCCTCTGCAGAAAGTGCCTCCGCCCACTGTCGAG GGTCGAGACCAAGTGCCTCGGAGCCCTGCGGTTCAATGCCGTGTGCTGCTGCCCGCTG GCCGTCCCGCTGCTCCTGCGACCGCGTCACCTGTGACGGGGCTACAGGCCCGGTG TGCCCAAGGACGGGCGCACGTATGACAGTGATTGCTGGCGGACGAGGCTGAGTGCCG CAGCAGCGTGCCATCCCCAGCAAGCACCAGGCGCCGTGTGACGAGCCCGCTCCCAT GCCCTCGGGGTGAGTGTGATTGGGGCGACGCTGTGCTGTGAAGAACGGGCGAGCAGC GTGTGAATGCTGCAGGCGTGTGAGCCTCTACGATCCTGTGTGCGGCGAGCAGCGC GTACATACGGCAGCGGTGCGAGCTGGAGGCCACGGCTGTACCCTCGGCGGGGAGA TCCAGGTGGCGCGCAAGGACCCCTGTGACCGCTGCGGGCAGTCCGCTTTGGAGCCCT		

GTGCGAGGCCGAGACCGGGCGCTGCGTGTGCCCCCTCTGAATGCGTGGCTTTGGCCACG
 CCCGTGTGTGGCTCCGACGGGCACACGTACCCACGAGTGCACTGCTGCACGTGCACG
 CCTGCACACACAGATCAGCCTGCACGTGGCCTCAGCTGGACCCCTGTGAGACCTGTGG
 AGATGCCGTGTGTGCTTTTGGGGCTGTGTGCTCCGACGGGCAGTGTGTGTGCTCCCGG
 TGTGAGCACCCCCCGCCCGGCCCGTGTGTGGCAGCGACGGTGTCACTACGGCAGTG
 CCTGCGAGCTACGGGAAGCCGCTGCCTCCAGCAGACACAGATCGAGGAGGCCCGGGG
 AGGGCCGTGCGAGAGCCGAGTGCCTTCCGGAGGCTTGGCTTCTGGGGAGGACGGT
 GACTGTGAGCAGGAGCTGTGCCGGCAGCGCGGTGGCATCTGGGACGAGGACTCGGAGG
 ACGGGCCGTGTGTCTGTGACTTCAGCTGCCAGAGTGTCCAGGACAGCCCGGTGTGCGG
 CTCAGATGGGGTCACTACAGCACCGAGTGTGAGCTGAAGAAGGCCAGGTGTGAGTCA
 CAGCGAGGGCTCTACGTAGCGGCCAGGGAGCCTGCCAGGCCCCGCTTCCGCCCGC
 TGCCGCTGTGGCCCCCTTACACTGTGCCCAGACGCCCTACGGCTGCTGCCAGGACAA
 TATCACCGCAGCCCGGGCGTGGGCTGGCTGGCTGCCCCAGTGCCTGCCAGTGAAC
 CCCATGGCTCTTACGGCGGCACCTGTGACCAGCCACAGGCCAGTGTCTTCCGCCG
 CAGGTGTGGGGGCTCAGGTGTGACCGCTGTGAGCCTGGCTTCTGGAACTTTCGAGG
 CATCGTACCGATGGCCGGAGTGGCTGTACACCTGCAGCTGTGATCCCCAAGGCGCC
 GTGCGGGATGACTGTGAGCAGATGACGGGGCTGTGCTCGTGTAGCCCCGGGGTGGCTG
 GACCCAAGTGTGGGAGTGTCCAGACGGCCGTGCCCTGGGGCCCGCGGGCTGTGAAGC
 TGACGCTTCTGCGCTGCGACCTGTGCGGAGATGCGCTGTGAGTTCGGTGTGCGGGTGC
 GTGGAGGAGTCTGGCTCAGCCCACTGTGTGTGCCCGATGTCTACCTGTCCAGAGGCCA
 ACGCTACCAAGGTCTGTGGGTGAGTGGAGTACATACGGCAACGAGTGTGAGTGA
 GACCATCGCTGCGCCAGGGCTGCAAACTCTATCCAGAGCCTGGGGCCCGTGGCAG
 GAGGCTGTGTCTCCAGCACTACCCGACATCTGCCCTCCGTGACTGTGACCACCCAG
 GGCTCCTCTGAGCCAGGCACTGCCGGCCCCCCCCGGCGCCCTCCCCCTGGCTCCAG
 CAGTACCGCACAGCCAGACACCCCTCCGCCCTCATCGGCACTCGGACCACTGCC
 AGCGTCCCGAGGACACCGTGTGGCCGTGTGACGGTGGCCCCCAGCGACCTCCCTCC
 CTGCAACCGAGCCTGGTGGCGTCCGCCCTTGGTGAATCTGCGAGCACTGTGGAAGCAG
 CGATGAGGAACTGAGCGGGGACAGGAGGCCAGTGGGGGTGGCTGTGGGGGGCTCGAG
 CCCTTGGAGGGCAGCAGCGTGGCCACCCCTGGGCCACCTGTGCGAGAGGCTTCTGTCT
 ACAACTCCGCGTGTGGGTGTGCTCTGATGGGAAGACGCCCTCGCTGGACGCGAGGG
 CTCCAACCTGCCCGCCACCAAGGTGTTCAGGGCGTCTGGAGCTGGAGGGCGTCCGAG
 GGCCAGGAGCTGTTCTACAGCCCGAGATGGCTGACCCCAAGTCAGAACTGTTCCGGG
 AGACAGCCAGGACCACTGAGAGCACTTGGACGACCTTCTCCGGAATTCAGACCTCAA
 GAAGGATTCCGGAGTGTCCGCTTGGCGGACCTGGGGCCCGGCAAACTCCGTCCCGGCC
 ATTGTGGATGTGCACTTGTACCCACACAGCCTTCAGGGCACCAGCGTGGCCCGGG
 CCCTGTCTCGGCAGATCCAGGTGTCCAGGCGCGGTCTTGGGGCGGAGGCGGCCGCT
 GCAGGAGCAGTGCATTTATGGAATTTGACTGGTTTCTGCGTTTATCACGGGGGCC
 ACGTCAGGAGCATTGTGCGGGAGCCACGGCCAGAGCCACCACTGCACTCGCGCTGCG
 CGTCTCTGCTGTGACCCCTCGGGCCCCGACCCCACTGACACAAGCCAGCCCGTGG
 CAAGACCGGAGCCCGCCACCACTGCGCCCCCACCCTGCCCCCAGCCGCTGTG
 CCCGACGTGCGCCCCCGGCCCGCCAGCAGCCTCCAAGCCCTGTGACTCACAGCCCT
 GCTTCCAGGGGGGACCTGCCAGGACTGGGCATTGGGCGGGGGTTCACCTGCACTG
 CCCGGCAGGAGGGGAGCGCGTCTGTGAGAAGGTGCTTGGCGCCCTGTGCGCGCC
 TTCGAGGGCCGCTCCTTCTGGCCTTCCCCACCCTCGCGCCTACACAGCTGCGCC
 TGGCACTGGAATTCGGGCGCTGGAGCCTCAGGGGCTGTGCTGTACAATGGCAACGC
 CCGGGGCAAGGACTTCTGGCATTGGCGCTGCTAGATGGCCGCTGCGAGCTCAGGTTT
 GACACAGGTTTCGGGGCCGCGGTGCTGACAGTGCCGTGCCGTGAGGAGCCGGGCGT
 GGCACCGCTGGAGCTGTCCCGCACTGGCGCGGGGACCCCTCTCGTGGATGGTGA
 GACCCCTGTCTGGGCGAGAGTCCAGTGGCACCAGCGGCTCAACCTGGACACAGAC
 CTCTTGTGGGCGCGTACCCGAGGACAGGCTGCCGTGGCGCTGGAGCGGACCTTCG
 TGGGCGCCGCGCTGAGGGGGTGCATCCGTTTGTCTGGAGCTCAACACACAGCGCTGA
 GCTTGGCATTGGGCGGGGGTGCACCCGAGGCTCTGGCGTGGGCGAGTGCAGGGGAC
 CACCCCTGCTGCCCAACCCCTGCCATGGCGGGGCCCATGCCAGAACCTGGAGGCTG
 GAAGGTTCCATTGCGCAGTGCCCGCCCGCGCGCTCGGACCACTGTGCCGATGAGAA
 GAGCCCTGCCAGCCCAACCCCTGCCATGGGGCGGCGCCCTGCCGTGTGCTGCCCGAG
 GGTGTGTCTCAGTGCAGTGCCTTGGGGCGTGGGGCACCTTCTGCCAGACAGCCT
 CGGGGCGAGAGGCTCTGGGCCCTTCTGGCTGACTTCAACGGCTTCTCCACCTGGA
 GCTGAGAGGCTGCAACCACTTGCACGGGACCTGGGGGAGAGATGGCGCTGGAGGCC
 GTGTTCTGGCACGAGGGCCAGCGGCTCTGTCTTACAACGGGCGAGAAGACGGACG
 GCAAGGGGGACTTGTGTGCTGGCACTGCGGGACCGCGCTGGAGTTCCGCTACGA
 CCTGGGCAAGGGGCGAGCGGTATCAGGAGCAGGAGCCAGTCAACCTGGGAGCTTGG
 ACCAGGCTTCACTGGAGCGAAGCGGCGCAAGGGTGCCTTGGTGTGGGCGAGCGCC
 CCCGTGTGTGGGGAGTCCCCGAAATCCCGCAAGGTTCCGCACACCGTCTCAACCT
 GAAGGAGCCGCTTACGTAGGGGGCGCTCCCGACTTCAGCAAGCTGGCCGCTGTGCT
 GCCGTGTCTGTGCTTGCAGCGTGCATCCAGCTGTGCTTCTCCGAGGCGCCAGC
 TGCTGACCCCGGAGCAGTGTGCGGCGAGTGGACGTACGTCCTTTCAGGTCACCC
 CTGCAACCGGGCCTCAGGCCACCCCTGCTCAATGGGGCTCTGCGTCCCGAGGGG
 GCTGCTATGTGTGCTGTGCCGGGGATTCTCAGGACCGCACTGCGAGAAGGGG
 TGGTGGAGAAGTCAAGGGGGAGTGGATACCTTGGCTTTCAGGGGCGGACCTTGT
 CGAGTACCTCAACGCTGTGACCGAGAGCGAGAAGGCACTGCAGAGCAACCACTTTGAA
 CTGAGCCTGCGCACTGAGGCCACGAGGGGCTGGTGTCTGGAGTGGCAGGGCACGG
 AGCGGGCAGACTATGTGGCACTGGCCATTGTGGAAGGGCACTGCAACTGAGCTACAA
 CCTGGGCTCCAGCCCGTGGTGTGCGTTCACCGTGGCGTCAACACCAACCGCTGG

	TTGCGGGTCGTGGCACATAGGGAGCAGAGGGAAGGTTCCCTGCAGGTGGGCAATGAGG CCCCCTGTACCGGCTCCTCCCGCTGGGCGCCACGCAGCTGGACACTGATGGAGCCCT GTGGCTTGGTGAGTGTTTTGGGGAGACTAGAGAGGGATGCCCAAGGGTCTCAGATATC CGAGGGACAGACTCCACCCCCAGCGCCCCACCTTGAGTCAGGGTGCATGTGAGCCGG CGGGCTGGGCTCTCTTCTCCGCTGTAGCCCCCTGCAGTTCAGTCTGTGGGGCCGG GAGGCGGGTGCCAGGTGTGGGCCCCCTGCTGGTCACCTGCTCTGTGGGGTGCCCATC AGCATCACTGAGTCACAGCCGGGTGACTCCCAGTGTCTGTGCTGCAGGGGCTGCCGG AGCTGCCCGTGGGCCCAGCACTGCCAAGGCCTACGGCACAGGCTTTGTGGGCTGCTT GCGGGATGTGGTGGGCGCGCACCCGCTGCACCTGCTGGAGGACGCGCTCACCAG CCAGAGCTGCGGCCCTGCCCCACCCATGAGCTGGCACCAGAGCCCCGCGCCGCT		
	ORF Start: ATG at 37	ORF Stop: TGA at 6466	
	SEQ ID NO: 106	2143 aa	MW at 225054.6kD
NOV32f, CG94946-06 Protein Sequence	MRHGRPVPPGPAAGRPLPLLVAAACVLPAGGTCPERALERREEEANVLTGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLLDGGNKVVISGFGDPLICDNQVSTGDT RIFFVNPAPPYLWPAHKNELMLNSSLMRITLRLNEEVEFCVEDKPGHTFTVPPTPPD ACRGMLCGFGAVCEPNAEGPGRASCVCCKSPCSVVAVPVCSDASTYSNECELQRAQC SQORRIRLLSRGPCSRDPCSNVTCSEGSTCARSADGLTASCLCPATCRGAPEGTVCG SDGADYPGECQLLRRACARQENVFKFDGPDPCQALPDPSRSCRVPNRRPEMLL RPESCPARQAPVCGDDGVTYENDCVMGRSGAARGLLLQKVRSGQCQGRDQCEPCRFN AVCLSRGRPRCSCDRVTCGAYRPVCAQDGRYDSDCWRRQAECRQRAIPSKHQGP CDQAPSPCLGVQCAFAGATCAVKNQQAACECLQACSSLYDPVCGSDGVITYGSACELEAT ACTLGREIQVARKGPCDRCGQCRFGALCEAETGRCVCPSECVALAQPVCSDGHTYPS ECMLHVHACTHQISLHVASAGPCETCGDAVCAFGAVCSAGQVCPRCEHPPPGVPVCGS DGVITYGSACELEAELQOTQIEARAGPCEQAECGSGSGSGEDGDCEQLCRQGG IWDEDESDGPCVDFSCQSVPGSPVCGSDGVITYSTECLEKKARCESQRLGYVAAQAGAC RGPAPAPLPFVAPLHCAQTPYGCCQDNITAARGVGLAGCPACQCNPHGSGYGTCDPA TGQCSRCRPGVGLRCDRCPEGFNFRGIVTDGRSGCTPCSCDPQGAVRDDCEQMTGLC SCKPGVAGPKCGQCPDGRALGPAGCEADASAPATCAEMRCFEGARCVEESGSAHCVC MLTCPEANATKVCSDGVITYGNECQLKTIACRQGLQISIQSLGPCQEAAPSTHTPTSA SVTVTTPGLLLSQALPAPPALPLAPSSTAHSQTTTPPSSRPRTTASVPTTVPVLT VPPTAPSPAPSLVASAFGESGSTDGSSDEELSGDQASGGSGGLEPLEGSSVATPGP PVERASCYNALGCCSDGKTPSLDAEGSNCPATKVFQGVLELEGVEGQELFYTPEMAD PKSELFGETARSIESTLDDLFRNSDVKKDFRSVRLRDLGPGKSVRAIVDVHFDPTTAF RAPDVARALLRQIQVRRRSLGVRRPLQEHVRFMDFWFFAFITGATSGAIAAGATAR ATTASRLPSSAVTPRAPHPSHTSQPVAKTTAAPTTRRPPTTAPSRVPGRRPPAQQPP KPCDSQPCFHGGTCQDVALGGGFTCSCPAGRGGAUCEKVLGAPVPAFEGRSFLAFPTL RAYHTLRLEALEFRALPQGLLLYNGNARGKDFLALALLDGRVQLRFDTGSGPAVLTS VPVEPGQWHRLELSRHWRGTLSDVGETPVLGESPSGTDGLNLDLDFVGGVPEDQAA VALERTFVAGLRGCIRLLDVNNQRLLELIGPGAATRGSGVGECDHPLCPNPCHGGA PCQNLAEGRFHCQPPGRVGPCTCADEKSPCQPNPCHGAAPCRVLEGGGAQCECPLGRE GTFCQTAGSQDGSFPFLADFNFSHLELRGLHTIARDLGEKMALEAVFLARGPSGLLL YNGQKTDGKDFVSLALRDRRLFRYDLGKGAIVRSREPTVLGAWTRVSLERNRKG ALRVGDGPRVLGESPKSRKVPHTVNLKEPLYVGGAPDFSKLARAAAVSSGFDGAIQL VSLGGRQLLTPEHVLQRQVDVTSFAGHPCTRASGHPCLNGASCVPREAAVCLCPGGFS GPHCEKGLVEKSAGDVTDLAFDGRTFVEYLNATVESEKALQSNHFLSLRTEATQGLV LWSGKATERADYVALAIVDGHQLSLYNLGSQPVVLRSTVPTVNRWLRVVAHREQREG SLQVGNEAPVTGSSPLGATQLDLDGALWLGECFGETREGCPRVSDIRGTDSTPQRPL SQGACEPAGWALFRCSPSSQCCGAGRRVPRCGPPAGHLLVGVPISTITESQPGDSHC LCCRGLPELPVGPALPKAYGTGFVGLCLRDVVVGRHPLHLEDVATKPELRPCPTP		
	SEQ ID NO: 107	5688 bp	
NOV32g, CG94946-07 DNA Sequence	CCGGCGCGGCCCCGCGCGCTCTTCCGCGCGCTCTCGCATGCGCCATGGCCGGCCGGTCC CACCCGGGCGCGCTGCGGGGCGGCGCTGCTGCTCTCCTTGTGGTGGCCGCGTGCCT CCTGCGCGGAGCGCGCGGACATGCCCGGAGCGCGCTGGAGCGGCGCGAGGAGGAG GCGAACGTGGTGCTCACCGGACGGTGGAGGAGATCCTCAACGTGGACCCGGTGACG ACACGTACTCTGCAAGGTTCCGGGCTGCGCGTACTTGAGGGGCAAGACCTGGTGGC CCGGGAGAGCCTGCTGGACGCGCGCAACAAGGTGGTGATCAGCGCTTTGGAGACCCC CTCATCTGTGACAAACAGGTGTCCACTGGGGACACCAAGGATCTTCTTTGTGAACCTG CACCCCCATACCTGTGGCCAGCCCAAGAAGAGCTGATGCTCAACTCCAGCCTCAT GCGGATCACCTGCGGAACCTGGAGGAGGTGGAGTTCTGTGTGGAAGATAAACCCGGG ACCCACTTCACTCCAGTGCCCTCCGACGCTCTGATGCGTGCAGGGGAATGCTGTGCG GCTTCGCGCGCGTGTGCGAGCCCAACGCGGAGGGGCGGGCGGGCGTCTGCGTCTG CAGAAGAAGCCCGTGCCCGAGCGTGGTGGCGCCTGTGTGGTGGTGGACGCGCTCCACC TACAGCAACGAATGCGAGCTGCAGCGGGCGCAGTGACGACGAGCGCGCATCCGCC TGCTCAGCCGCGGGCCGTGCGGCTCGCGGGACCCCTGCTCAACGTGACCTGCAGCTT CGGCAGCACCTGTGCGCGCTCGGCCGACGGGCTGACGGCTCTGTCCTGTGCCCCGCG ACCTCGGTGGCGCCCCGAGGGGACCGTCTGCGGCAGGACGCGCGCGCATCCCGG GCGAGTGCCAGCTCCTGCGCGCGCTGCGCCCCGAGGAGAATGTCTTCAAGAAGTT CGACGGCCCTTGTGACCCCTGTGAGGGCGCCCTCCTGACCCGAGCCGAGCTGCCGT GTGAACCCGCGCAGCGCGCCCTGAGATGCGCCTACGCGCCGAGAGTGCCTGCGCC GGCAGGCGCCAGTGTGTGGGACGACGAGTCACTACGAAACGACTGTGTCTGAGG		

CCGATCGGGGCGCCCGGGGTCTCTCTCGAGAAAGTGCCTCCGGCCAGTGCCAG
GGTCGAGACCAAGTGCCCGGAGCCCTGCCGGTTCAATGCCGTGTGCTGTCCCGCGTG
GCCGTCCCCGCTGCTCCTGCGACCGCGTCACTGTGACGGGGCTACAGGCCCGTGTG
TGCCACGAGACGGGCGCACGTATGACAGTGATTGCTGGCGGACAGCGCTGAGTGCCGG
CAGCAGCGTGCCATCCCCAGCAAGCACCAGGGCCCGTGTGACAGGCCCGTCCCCAT
GCCTCGGGGTGCACTGTGCAATTTGGGGCGACGTGTGCTGTGAAGAACGGGACAGCAGC
GTGTGAATGCCTGCAGGCGTGTGCTGAGCCTCTACGATCCTGTGTGCGGACGACGGC
GTCACATACGGCAGCGCTGCGAGCTGGAGGCCACGGCCTGTACCTCGGGCGGGAGA
TCCAGGTGGCGCGCAAAGACCTGTGACCGTGCAGGCGAGTGCCGCTTTGGAGCCCT
GTGCGAGGCCGAGACCGGGCGCTGCGTGTGCCCTCTGAATGCGTGGCTTTGGCCAG
CCCGTGTGTGGCTCCGACGGGACACGTACCCAGCGAGTGACCTGTGCACGTGCACG
CCTGCACACACAGATCAGCCTGCACGTGGCCTCAGCTGGACCTGCGAGACCTGTGG
AGATGCCGTGTGTCTTTGGGGCTGTGTGCTCCGACGGGCGAGTGTGTGTGTCCCGG
TGTGAGCACCCCGCGCCCGCCGTGTGTGGCAGCGAGGTGTACCTACGGCAGTG
CCTGCGAGCTACGGGAAGCCCGCTGCTCCAGCAGACAGATCGAGGAGCCCGGGC
AGGGCGGTGCGAGCAGGCGAGTGCGGTTCCGGAGGCTGTGGCTGTGGGAGGACGGT
GACTGTGAGCAGGAGTGTGCCGGCAGCGCGTGGCATCTGGGACGAGGACTCGGAGG
ACGGGCGGTGTGTCTGTGACTTCACTGCTGCCAGAGTGTCCAGGCGAGCCCGTGTGCGG
CTCAGATGGGGTCACTACAGCACCGAGTGTGAGCTGAAGAAGGCCAGGTGTGAGTCA
CAGCGAGGGCTCTACGTAGCGGCCAGGGAGCCTGCCAGGCCCCACCTTCGCCCCGC
TGCCGCTGTGGCCCCCTTACTGTGCCAGACGCCCTACGGTGTGCCAGGACAA
TATCACCGCAGCCCGGGCGTGGGCTGTGGCTGGCTGCCAGTGCCTGCCAGTGCAAC
CCCCATGGCTCTTACGGCGGCACCTGTGACCCAGCCACAGGCCAGTGCTCCTGCCGCC
CAGGTGTGGGGGCGCTCAGGTGTGACCGCTGTGAGCCTGGCTTCTGGAATTTTCGAGG
CATCGTCAACGATGGCCGAGTGGCTGTACACCTGACGTGTGATCCCCAAGGCGCC
GTGCGGATGACTGTGAGCAGATGACGGGGCTGTGCTCGTGTGAAGCCGGGTGTG
GACCAAGTGTGGGCGAGTGTCCAGACGGCGTGCCTGGGCCCCGGGGCTGTGAAGC
TGACGCTTCTGCGCCTGCGACCTGTGCGGAGATGCGCTGTGAGTTCCGTGCGCGGTGC
GTGAGGAGTCTGGCTCAGCCCACTGTGTCTGCCGATGCTCACTGTCCAGAGGCCA
ACGTACCAAGGTCTGTGGGTGAGTGGAGTACATACGGCAACGAGTGTGAGTGA
GACCATCGCTGCCAGCGGTGTCACTACGCCAGGGCCTGCAATCTCTATCCAGAGC
CTGGGCGCGTGCAGGAGGCTGTGCTCCAGCACTCACCGACATCTGCCTCCGTGA
CTGTGACACCCAGGGCTCCTCCTGAGCCAGGCACTGCCGGCCCCCGGGCGCCT
CCCCCTGGCTCCAGAGTACCGCACACAGCCAGACACCCCTCCGCCCTCATCGGA
CCTCGGACCACTGCCAGCGTCCCCAGGACCACTGTGGCGCGTGTGACGGTGGCCCC
CCACGGCACCTCCTCTGCAACCCAGCCTGGTGGCGTCCGCTTTGGTGAATCTGGCAG
CACTGATGGAAGCAGCGATGAGGAAGTGAAGCGGGACAGGAGGCGAGTGGGGGTGGC
TCTGGGGGGCGGAGCCCTTGGAGGGCAGCAGCGTGGCCACCCCTGGGCCACCTGTG
AGAGGGCTCCTGTACAACCCCTGCCATGGGGCGGCGCCTGCCGTGTGCTGCCGA
GGGTGTGTCTCAGTGCAGTGCCTCCTGGGGCGTGAAGGCACTTCTGCCAGACAGC
TCGGGGCAGGACGGCTCTGGGCCCTTCTGGCTGACTTCAACGGCTTCTCCACCTGG
AGCTGAGAGGCTGCACACCTTTGACAGGGACCTGGGGGAGAAGATGGCGCTGGAGGT
CGTGTCTCTGGCACGAGGCCAGCGGCTCTGCTCTACAACGGGCAGAACAGCGAC
GGCAAGGGGACTTCTGTGCTGTGGCACTGCGGGACCGCCCTGGAGTTCCGCTACG
ACCTGGGCAAGGGGGCAGCGGTCTCAGGAGCAGGGAGCAGTCACTGGGAGCCTG
GACCAAGGTCTCACTGGAGGAAACGGCGCAAGGGTGCCTTCGCTGTGGGCGAGCGC
CCCCGTGTGTGGGGGAGTCCCGGTTCCGCACACCGTCTCACTGAAGGAGCGCG
TCTACGTAGGGGGCGCTCCGACTTCAAGCAAGTGGCCCGTGTGCTGCCGTGTCTC
TGGCTTCGACGGTGCATCCAGCTGGTCTCCTCGGAGGCGCCAGCTGTGACCCCG
GAGCAGTGTGCGGCAAGTGGACGTCACTCTTTGACAGTCAACCTGACCCCGG
CCTCAGGCCACCCCTGCCTCAATGGGGCTCCTGCGTCCCGAGGGAGGCTGCTATGT
GTGCTGTGTCCCGGGGATTCTCAGGACCGCACTGCGAGAAGGGGCTGGTGGAGAAG
TCAGCGGGGACGTGGATACCTTGGCCTTTGACGGGCGGACCTTGTGAGTACCTCA
ACGTGTGACCGAGAGCGAGAAGGCACTGCAGAGCAACCACTTGAAGTGAAGCTGCG
CACTGAGGCCACGAGGGGCTGGTGTCTGAGTGGCAAGGCCACGGAGCGGGCAGAC
TATGTGGCACTGGCATTGTGGAGGGCACCTGCAACTGAGTACAACCTGGGCTCCC
AGCCCGTGGTGTGCTGCCCTCCACCGTGGCCGTCAACCAACCGTGGTTCGGGTGTG
GGCACATAGGGAGCAGAGGGAAGGTTCCCTGACGGTGGGCAATGAGGCCCTGTGACC
GGCTCCTCCCGCTGGGCGCCACGCACTGGACACTGATGGAGCCCTGTGGCTTGGGG
GCCTGCCGGAGTGCCTGTGGGCGGCACTGCCCAAGGCTACGGCACAGGCTTTGT
GGGCTGCTTGGGGATGTGGTGGTGGGCGGCAACCGCTGCACCTGTGGAGGAGCGC
GTCACCAAGCCAGAGTGCGGCCCTGCCCAACCCATGAGCTGGCACAGAGCCCCG
GCCCCGTGTAATATTTCTATTTTGTAACTTGTGTCTTTTGATATGATTTCTT
GCCTGAGTGTGGCGGAGGAGTGTGGCCCGGCTCCTTCCGTCCAGGACAGCGGT
GCTGCAGACAGACTAGTGTGAGGGATGGACAGGCGAGGTGGCAGCGTGGAGGGCTC
GGCGTGGATGGCAGCCTCAGGACACACACCCCTGCCTCAAGTGTGTGAGCCCGCCT
TGCACTGCGCCTGCCCAAGGTGTCCCGCGGGGAAGCAGCCCGGCTCCTGAATCAC
CCTGCTCCTGTCAGGCGGACTCGTGTCCAAAAGGAAGGGGCTGTGAGGTGTGAT
GGGGCCCTTCTCCGGGTGACCCACAGGGCCTTCCAGCCCTTATTGAGCTGTCTC
CTTCTGTGTGTCTGTGACCTGCTCGGCTCCTGCGCAATACTGTGACTTCCA
AACAATGTACTGTGGGACAGCTGTGCTGTCTCCGCTGTGCTGCGCGAGCCCC
AGGCTGTGAGGAGCAGAGGCCAGACAGGGCCGATCTGGGTGTCTGACCTCAGCT
GGCCCTGCCAGCCACCTGGACATGACCGTATCCCTTGGCACACCCAGGCCCTGC

	GAGGGGCTATCGAGAGGAGCTCACTGTGGGATGGGGTTGACCTCTGCCGCTGCCTGG GTATCTGGGCTGGCCATGGCTGTGTCTTTCATGTGTGATTTTATTTGACCCCTGGA GTGGTGGGTCTCATCTTTCCCATCTCGCCTGAGAGCGGCTGAGGGCTGCCTCACTGCA AATCCTCCCCACAGCGTCAGTGAAAGTCGTCCTTGTCTCAGAATGACCAGGGGCCAGC CAGTGTCTGACCAAGGTCAAGGGGCAGGTGCAGAGGTGGCAGGGATGGCTCCGAAGCC AGAA
--	---

	ORF Start: ATG at 37	ORF Stop: TGA at 4735	
	SEQ ID NO: 108	1566 aa	MW at 164102.4kD
NOV32g, CG94946-07 Protein Sequence	MRHGRPVPPGPAAGRPLPLLVVAACVLPAGAGGTCPERALERREEEANVVLGTVEEI LNVDPVQHTYSCKVRVWRYLKGKDLVARESLLDGNGKVVISGFGDPLICDNQVSTGDT RIFVFNAPPYLPWAHKNELMLNSSLMRITLRNLEEEVFCVEDKPGTHFTFPVPTPPD ACRGMLCGFVAVCEPNAEGPGRASCVCCKSPCSVVAVPCGSDASTYSNECELQRAQC SQRRIRLLSRGPGRSDPCSNVTCSTGSCARSADGLTASCLCPATCRGAPEGTVCG SDGADYPGECQLLRACARQENVFKKFDGPCDPCQALPDPSRSCRVPNPTRRPEMRL RPESCPARQAPVCGDDGVTYENDCMGRSGAARGLLQKVRSGQCQGRDQCEPECRFN AVCLSRGRPRCSDRVTCDAAYRPVCAQDGRTYSDCWQRQAECRQQRAPSKHQGP CDQAPSPCLGVQCAFATCAVKNQGAACECLQACSSLYDPVCGSDGVTYGSACELEAT ACTLGREIQVARKGPCDRCGQCRFGALCEAETGRCVCPSECVALAQPVCGSDGHTYPS ECMLHVHACTHQISLHVASAGPCETCGDAVCAFGAVCSAGQCVCPRCEHPPPGPVCGS DGVTYGSACELREAACLQQTQIEEARAGPCEQAECGSGSGSGEDGDCEQLCRQGG IWEDESDGPCVDFSCQSVPGSPVCGSDGVTYSTECCLKKARCESQRGLYVAAQGAC RGPTFAPLEPPVAPLHCAQTPYGCCQDNITAARGVGLAGCPSACQCNPHSGSYGGTCDPA TGQCSRCRPGVGLRCRCEPGFWNFRGIVTDGRSGCTPCSCDPQGAVERDDCEQMTGLC SCKPGVAGPKCGCPDGRALGPAGCEADASAPATCAEMRCEFGARCVEESGAHCVCP MLTCPEANATKVCSDGVTYGNECQLKTIACRRCHLRQGLQISIQSLGPCQEAAPST HPTASAVTTPGLLLSQALPAPPGALPLAPSSTAHSQTTPPPSSRRTTASVPRTTV WPVLTVPPTAPSPAPSLVASAFGESGSDTGSSDEELSGDQEAAGGGSGGPELEGSSV ATPGPPVERASCYNPCHGAAPCRVLEPEGGAQCECPLGREGTFCQTASQDGGSGPFLAD FNGFSLHLELRLHTFARDLGEKMALEVVFLARGPSGLLLYNGQKTDGKGDFVSLALRD RRLEFRYDLKGAAVIRSREPVTLAGWTRVSLERNRKGALRVGDGPRVLGESPVPH VLNLKEPLYVGGAPDFSKLARAASVSSGFDGAIQLVSLGGRQLLTPEHVLRVQDVTSF AGHPCTRASGHPCLNGASCVPREAAVYVCLCPGGFSGPHCEKGLVEKSAGDVTDLAFDG RTFVEYLNVAVTESEKALQSNHFELSLRTEATQGLVLWSGKATERADYVALAIVDGHLO LSYNLGSQPVVLRSTVPVNTNRWLRVVAHREQREGSLQVGNAPVTVGSSPLGATQLDT DGALWLGGLELPELVGPALPKAYGTGFVGLRDVVVGRHPLHLEDAVTKPELRPCPTP		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 32B.

Table 32B. Comparison of NOV32a against NOV32b through NOV32g.		
Protein Sequence	NOV32a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV32b	34..1536 34..1549	1072/1557 (68%) 1128/1557 (71%)
NOV32c	1913..2053 136..276	141/141 (100%) 141/141 (100%)
NOV32d	34..549 34..549	505/516 (97%) 505/516 (97%)
NOV32e	34..1536 34..1528	1170/1515 (77%) 1201/1515 (79%)
NOV32f	34..2002 34..2002	1758/1969 (89%) 1758/1969 (89%)
NOV32g	34..1536 34..1550	1081/1562 (69%) 1134/1562 (72%)

Further analysis of the NOV32a protein yielded the following properties shown in Table 32C.

Table 32C. Protein Sequence Properties NOV32a	
PSort analysis:	0.7618 probability located in outside; 0.1900 probability located in lysosome (lumen); 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 34 and 35

A search of the NOV32a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 32D.

Table 32D. Geneseq Results for NOV32a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV32a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAW26609	Human agrin - Homo sapiens, 492 aa. [WO9721811-A2, 19-JUN-1997]	1591..2053 22..492	460/471 (97%) 461/471 (97%)	0.0
AAB93754	Human protein sequence SEQ ID NO:13424 - Homo sapiens, 413 aa. [EP1074617-A2, 07-FEB-2001]	583..968 1..386	381/386 (98%) 384/386 (98%)	0.0
AAY73993	Human prostate tumor EST fragment derived protein #180 - Homo sapiens, 416 aa. [DE19820190-A1, 04-NOV-1999]	1634..2053 1..416	414/420 (98%) 414/420 (98%)	0.0
AAB31889	Amino acid sequence of a human protein - Homo sapiens, 4393 aa. [WO200105422-A2, 25-JAN-2001]	1355..2052 3639..4393	252/794 (31%) 352/794 (43%)	7e-88
ABG23265	Novel human diagnostic protein #23256 - Homo sapiens, 4436 aa. [WO200175067-A2, 11-OCT-2001]	1355..2051 3672..4435	252/803 (31%) 350/803 (43%)	7e-85

In a BLAST search of public sequence databases, the NOV32a protein was found to have homology to the proteins shown in the BLASTP data in Table 32E.

Table 32E. Public BLASTP Results for NOV32a				
Protein Accession Number	Protein/Organism/Length	NOV32a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
O00468	AGRIN PRECURSOR - Homo sapiens (Human), 2026 aa (fragment).	24..2053 1..2026	2022/2030 (99%) 2022/2030 (99%)	0.0
P25304	Agrin precursor - Rattus norvegicus (Rat), 1959 aa.	160..2053 51..1959	1558/1914 (81%) 1663/1914 (86%)	0.0
P31696	Agrin precursor - Gallus gallus (Chicken), 1955 aa.	128..2050 1..1952	1234/1970 (62%) 1479/1970 (74%)	0.0
Q90404	Agrin - Discopyge ommata (Electric ray), 1328 aa (fragment).	716..2051 1..1325	733/1353 (54%) 932/1353 (68%)	0.0
Q96IC1	UNKNOWN (PROTEIN FOR IMAGE:3544662) - Homo sapiens (Human), 488 aa (fragment).	1562..2053 1..488	486/492 (98%) 486/492 (98%)	0.0

PFam analysis predicts that the NOV32a protein contains the domains shown in the Table 32F.

Table 32F. Domain Analysis of NOV32a			
Pfam Domain	NOV32a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Kazal	201..246	25/61 (41%) 36/61 (59%)	7.2e-18
Kazal	276..321	21/62 (34%) 33/62 (53%)	5.1e-13
Kazal	346..393	18/61 (30%) 33/61 (54%)	7.9e-12
Kazal	420..465	21/61 (34%) 38/61 (62%)	4.1e-16
Kazal	494..538	24/61 (39%) 38/61 (62%)	3.6e-19
Kazal	559..603	19/61 (31%) 38/61 (62%)	1.5e-18
Kazal	624..668	26/62 (42%) 37/62 (60%)	1.5e-17
Kazal	709..754	24/62 (39%) 40/62 (65%)	1.2e-16
laminin_EGF	797..848	28/61 (46%) 46/61 (75%)	1.2e-20
laminin_EGF	851..895	21/59 (36%) 37/59 (63%)	4e-11
Kazal	927..973	25/62 (40%) 41/62 (66%)	5.3e-18
SEA	1134..1256	39/132 (30%) 112/132 (85%)	1.4e-36
EGF	1337..1370	16/47 (34%) 24/47 (51%)	0.00054
laminin_G	1404..1535	70/162 (43%) 119/162 (73%)	3.1e-53
EGF	1557..1589	16/47 (34%) 27/47 (57%)	5.1e-06
EGF	1596..1628	16/47 (34%) 25/47 (53%)	0.0002
laminin_G	1672..1807	70/161 (43%) 123/161 (76%)	5.1e-51
EGF	1826..1860	14/47 (30%)	2.3e-06

		25/47 (53%)	
laminin_G	1905..2036	59/161 (37%) 125/161 (78%)	1.7e-50

Example 33.

The NOV33 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 33A.

Table 33A. NOV33 Sequence Analysis		
	SEQ ID NO: 109	3354 bp
NOV33a, CG95165-01 DNA Sequence	<p>TGTGGCAGGAGCGATGCGGCGCCGCGCTACCTGCGGGACCGCTCCGAGGAGGCGGC GGGCGGCGGAGACGGGCTGCCGCGTCCCGGGACTGGCTCTACGAGTCTACTACTGC ATGAGCCAGCAGCACCCGCTCATCGTCTTCTGCTGCTCATCGTCATGGGCTCCTGCC TCGCCCTGCTCGCCGCTTCTTCGCGCTCGGCCTGGAAGTTGAAGACCATGTGGCGTT TCTAATAACAGTTCCAACTGCCCTGGCGATTTCTTTGGGATATTTATCCTGGTCTGC ATCGAGTCTGTGTTTAAAGAGCTGCTGCCGCTCTTCTCGTTGGTGATATGGATATGCC TTGTTGCCATGGGATACCTGTTTCATGTGTTTGGAGGCACCGTCTCTCCCTGGGACCA GGTATCGTCTTCTCTTCATCATCTTCGTGGTGTACACCATGCTGCCCTTCAACATG CGAGACGCCATCATTGCCAGCGTCTCACCTCCTCTCCACACCATCGTGCTAAGCG TCTGCCGTCTGCAACACCGGGAGGCAAGGAGCACCTGGTCTGGCAGATCCTGGCCAA TGTGATCATTTTCATCTGTGGGAACCTGGCGGGAGCCTACCATAAGCACTCATGGAA CTCGCTCTTCAGCAACATATCAGGACACCTGTAATTGCATCAAGTCGCGGATCAAGT TGGAAATTGAAAAACGTCAACAGGAGCGGCTTCTGCTCTCCCTGCTGCCGGCCACAT CGCCATGGAGATGAAGCGGAGATCATCCAGAGGCTGCAGGGCCCCAAGGCGGGCCAG ATGGAGAACACAAATAACTTCCACAACCTGTATGTGAAGCGGCATACAAAGTGAACA TCTTATACGCTGACATCGTTGGCTTTACCGGCTGGCAAGTGACTGCTCCCGGGAGA ACTAGTCCACATGCTGAATGAGCTCTTTGGAAAGTTTGTACAAATTGCAAGGAGAA GAATGCATGAGAATTAATAATTTAGGAGACTGCTACTACTGTGTATCTGGACTCCCTA TATCTCTCCCTAACCATGCCAAGAACTGTGTGAAAATGGGGCTGGACATGTGTGAAGC CATAAAGAAAGTGAGGGATGCTACTGGAGTTGATATCAACATGCGCGCTGGGCGTGCAT TCTGGGAATGTCTGTGTGGCGTGATTGGTCTGCAGAAGTGGAATATGATGTGTGGT CACATGATGTGACCTTGGCCAACCATGGAAGCTGGAGGGGTCCCTGGACGTGTTCAT CATTCTTCTGTACCTTGGAGCACTTGAATGGCGCTTATAAAGTGAGGAGGGAGAT GGTGACATTAGGGACCCATATTTAAACAGCACCTGGTGAAAACCTACTTTGTGATCA ACCCCAAGGGAGAACGACGAGCCCCAGCATCTCTTCAGACCTCGCCACACCCCTTGA TGGAGCCAAATGAGGGCCTCGGTCCGATGACCCGGTACTTGGAGTCCCTGGGGGGCA GCCAAGCCCTTTGCACACCTACATCACAGGGACAGCATGACACAGAGAACGGCAAGA TCAGCACACCGGATGTACCATGGGTGAGCATAATTTTCAAATCGCACCTTAAGAAC CAAGTCACAAAGAGAGATTTGAAGAAGAAATGAATGAAAGGATGATTAAGCAATT GATGGGATTAATGCACAGAAGCAATGGCTCAAGTCTGAAGACATTAGAGAATCTCAC TGCTTTTCTATAAAGTACTAGAAAAGAGTACCGGGCCACGGCACTGCCAGCGTT CAAGTATTATGTGACTTGTGCTGTCTCATATTCTTCTGCATCTTCATTGTGCAGATT CTCGTCTGCCCCAACGCTGTCTCTGGGCATCTCTTGGGGCTGCGTTCTCTTGTG TGGCCTTCATCTCTCTGCTGCTTTGCTGGACAGCTTCTGCAATGCAGCAAAAAGC CTCTCCCTGCTCATGTGGCTTTTGAAGTCTCGGGCATCATTTGCCAACCGCCCTGG CCACGGATCTCTCTACGATCATCACCAGCCATCATATTATGATGGCGGTGTTCAT ACATGTTTTCTGAGTGACTCAGAGGAACAAATCCCTCCAATGCCAACACAAACAA CACAACTTTTTCAGCTCAATAATCAGGTGGCGATTCTGCGTGGCAGAAATTTATTT TTCCTCCGTAATTTATCTACAGCTGCATTCTGGGACTGATATCCTGTTCCGTGTTC TGCGGGTAACTATGAGCTGAAGATGTTGATCATGATGGTGGCCTTGGTGGGCTACAA CACCATCTACTCCACACCCACGCCCAGTCTGGGGGACTACAGCCAGGTCTTATTT GAGAGACCAGGCATTGGAAAGACCTGAAGACCATGGGCTCTGTGCTCTCTCTATAT TCTTCATCAGCTGCTTGTCTGGGTAGACAGAATGAATATTACTGTAGGTTAGACTT CTTATGGAAGAACAAATCAAAAAGAGCGGGAGGAGATAGAGACCATGGAGAACCTG AACCGCTGTCTGGAGAACGTGCTTCCCGCGCACGTGGCTGAGCACTTCTGGCCA GGAGCCTGAAGAATGAGGAGCTATACCACAGTCTTATGACTGCGTCTGTGTCTGTT TGCCTCCATTCCGGATTTCAAAGAATTTATACAGAATCCGACGTGAACAGGAGGGC TTGGAATGCCTTCGGCTCCTGAACGAGATCATCGCTGACTTTGATGATCTTCTTTCCA AGCCAAAATTCAGTGGAGTTGAAAAGATTAAGACCATTTGGCAGCATATATGGCAGC AACAGGTCTGAGCGCTGTGCCAGCCAGGAGCACTCCAGGAGCCCGAGCGGCAGTAC ATGCACATTGGCACCATGGTGGAGTTTGTCTTTGCCCTGGTAGGGAAGCTGGATGCCA TCAACAAGCACTCTTCAACGACTTCAAATTGCGAGTGGGTATTAACCATGGACCTGT GATAGCTGGTGTGATTGGAGCTCAGAAGCCACAATATGATATCTGGGGCAACACTGTC</p>	

	AATGTGGCCAGTAGGATGGACAGCACCGGAGTCTCTGGACAAAATACAGGTTACCGAGG AGACGAGCCTCGTCCTGCAGACCTCGGATACACGTGCACCTGTGAGGAATAATCAA CGTGAAAGGAAAGGGGACCTGAAGACGTACTTTGTAAACACAGAAATGTCAAGGTCC CTTTCCAGAGCAACGTGGCATCTGAAGAGTCACCTTCATTTTGGCAAGAAGACTGT ATTTTCAGGAAGGTATCACACACTTTCTGACTGCAACTTCTGTCCCTT
--	--

	ORF Start: ATG at 15	ORF Stop: TGA at 3273	
	SEQ ID NO: 110	1086 aa	MW at 122956.2kD
NOV33a, CG95165-01 Protein Sequence	MRRRRYLDRSEEAGGGDGLPRSRDWLYESYYCMSQQHPLIVFLLIIVMGSCALLA VFFALGLEVEDHVAFLITVPTALAIFFAIFILVCIESVFKLLRLFSLVIWICLVAMG YLFMCFGGTVPWDQVSFFLFII FVVYTM L PFMNRDAIIASVLTSSSHITVLVSLCSA TPGGKEHLVWQILANV I I FICGNLAGAYHKHLMELALQQTYQDTCNCIKSRIKLEFEK RQOERLLLSLLPAHIAMEMKAEI IORLQGP KAGQMENTNNFHNLYVKRHTNVSIYAD IVGFTRLASDCSPGELVHMLNELFGKFDQIAKENECMRILKLGDCYYCVSGLPISLPN HAKNCVKMGLDMCEAIKKVRDATGVDINMRVGVHSGNVLCGVIGLQKWQYDVWSHDVT LANHMEAGGVPRVHISSVTLEHLNGAYKVEEGDGDIRDYPYLKQHLVKTYFVINPKGE RRSPQHLFRPRHTLDGAKMRASVRMTRYLESWGAAKPFAHLHHRDSMTTENGKISTTD VPMGQHNFNQRTLRTKSQKKRFEELNERM IQAIDGINAQKWLKSEDIQRISLLFYN KVLEKEYRATALPAFKYVVTCACLIFFCIFTVQILVLPKTSVLGISFGAAFLLLAFIL FVCFAGQLQCCKASPLLMWLLKSSGIIANRPWPRI SLTIITTAIILMMAVNMFFL SDSEETIPPTANTNTSFSASNQVAILRAQNLFFLPYFIYSCILGLISCSVFLRVNY ELKMLIMMVALVGYNTIILHTHAHVLGDYSQVLFERPGIWKDLKTMGSVSLSIFFITL LVLGRQNEYCRDLFLWKNKFKKEREETMENLNRVLENVLPAAHVAEHFLARSLKN EELYHQSYDCVCMFASIPDFKEFYTESDVNKEGLECLRLNEI IADFDLLSKPKFS GVEKIKTIGSTYMAATGLSAVPSQHSQEPERQYMHIGTMVEFAPALVGKLDAINKFS FNDFKL RVGINHG PVIAGVIGA QKPQYDIWGNTVNVASRMDSTGVLDKIQVTEETSLV LQTLGYTCTCRGI INVKGKGLKTYFVNTEMSRSLSQSNVAS		

Further analysis of the NOV33a protein yielded the following properties shown in Table 33B.

Table 33B. Protein Sequence Properties NOV33a	
PSort analysis:	0.8000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 65 and 66

A search of the NOV33a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 33C.

Table 33C. Geneseq Results for NOV33a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV33a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE0293 8	Human adenylate cyclase 25678 - Homo sapiens, 1086 aa. [WO200144453-A1, 21-JUN-2001]	1..1086 1..1086	1086/1086 (100%) 1086/1086 (100%)	0.0
AAB0200 6	Adenylyl cyclase type II-C2 C2 alpha domain - Homo sapiens, 1090 aa. [US6107076-A, 22-AUG-2000]	1..1086 1..1090	1039/1090 (95%) 1062/1090 (97%)	0.0
AAR9456 0	Rat adenylyl cyclase - Rattus sp, 1090 aa. [WO9608260-A1, 21-MAR-1996]	1..1086 1..1090	1039/1090 (95%) 1062/1090 (97%)	0.0
AAU0192 4	Human adenylate cyclase polypeptide - Homo sapiens, 1077 aa. [WO200125448-A1, 12-APR-2001]	22..1078 10..1069	605/1070 (56%) 772/1070 (71%)	0.0
AAB0200 8	Type IV adenylyl cyclase - Homo sapiens, 1064 aa. [US6107076-A, 22-AUG-2000]	22..1082 10..1062	609/1072 (56%) 769/1072 (70%)	0.0

In a BLAST search of public sequence databases, the NOV33a protein was found to have homology to the proteins shown in the BLASTP data in Table 33D.

Table 33D. Public BLASTP Results for NOV33a				
Protein Accession Number	Protein/Organism/Length	NOV33a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P26769	Adenylate cyclase, type II (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Rattus norvegicus (Rat), 1090 aa.	1..1086 1..1090	1039/1090 (95%) 1062/1090 (97%)	0.0
Q08462	Adenylate cyclase, type II (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Homo sapiens (Human), 887 aa (fragment).	200..1086 1..887	887/887 (100%) 887/887 (100%)	0.0
Q91WF3	SIMILAR TO ADENYLYL CYCLASE 4 (ADENYLYL CYCLASE TYPE 4) (EC 4.6.1.1) - Mus musculus (Mouse), 1077 aa.	22..1082 10..1075	612/1074 (56%) 780/1074 (71%)	0.0
CAC37757	SEQUENCE 2 FROM PATENT WO0125448 - Homo sapiens (Human), 1077 aa.	22..1078 10..1069	605/1070 (56%) 772/1070 (71%)	0.0
P26770	Adenylate cyclase, type IV (EC 4.6.1.1) (ATP pyrophosphate-lyase) (Adenylyl cyclase) - Rattus norvegicus (Rat), 1064 aa.	22..1082 10..1062	609/1072 (56%) 769/1072 (70%)	0.0

PFam analysis predicts that the NOV33a protein contains the domains shown in the Table 33E.

Table 33E. Domain Analysis of NOV33a			
Pfam Domain	NOV33a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Sodcu	388..395	6/8 (75%) 8/8 (100%)	0.73
guanylate_cyc	276..460	71/226 (31%) 151/226 (67%)	2.7e-68
guanylate_cyc	873..1073	89/227 (39%) 186/227 (82%)	2.1e-96

Example 34.

- 5 The NOV34 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 34A.

Table 34A. NOV34 Sequence Analysis

	SEQ ID NO: 111	3117 bp
NOV34a, CG95175-01 DNA Sequence	<p> CCTCCCCAGTAGCTGGGACTATGGGAGCGTGCCACCATGCCTGGTTAATTTTGTATT TTTAGTAGAGATGGGGTTTACCATGTTGGCCAGGCTTGTCTTCCCTCTCCTTAGTT ATCCTCCTGGATTCCAAAGCCTCCAGGCCGAGCTGGGCTGGACTGCACTGCCAAGTA ATGGGTGGGAGGAGATCAGCGGCGTGGATGAACACGACCGTCCCATCCGACGTACCA AGTGTGCAATGTGTGGAGCCCAACCAGGACAATGGCTGCAGACTGGGTGGATAAGC CGTGGCCGCGGGCAGCGCATCTTCGTGGAAGTGCAGTTACACTCCGTGACTGCAGCA GCATCCCTGGCGCCGCGGTACCTGCAAGGAGACCTTCAACGTCTACTACCTGGAAAC TGAGCCGACCTGGGCGTGGGCGTCCCGCCTAGGCGGCAAAATCGACACGATCGCG GCGGACGAGAGCTTACGACAGGGCGACCTGGGTGAGCGCAAGATGAAGTTGAACACAG AGGTGCGGAGATCGGACCGCTCAGCCGGCGGGGTTTCCACCTGGCCTTTCAGGACGT GGGCGCATGCGTGGCGCTTGTCTCGGTGCGCGTCTACTACAAGCAGTGCCGCGCCACC GTGCGGGGCTGGCCACGTTCCAGCCACCGCAGCCGAGAGCGCCTTCTCCACACTGG TGGAAAGTGGCCGGAACGTGCGTGGCGCACTCGGAAGGGGAGCCTGGCAGCCCCCAGC CATGCACTGCGGCGCCGACGGCGAGTGGCTGGTGCTGTGGGCGCTGCAGCTGCAGC GCGGATTCCAGGAGCGTGGTGAATCTGCGAAGGTATCTGTCCCCAGGGTTTACAC AGGTGTCCCGCGGCGGCCCTCTGCTCACCGTGCCAGAGCACAGCCGGGCCCTGGA AAACGCCTCCACCTTCTGCGTGTGCCAGGACAGCTATGCGCGCTCACCACCGACCCG CCCTCGGCTTCTGACACCGTCCGCCGTGCGCGCCGCGGAGCTGCACTGACAGCTGA GCCGCTCGCGCTGGTGTGCGACTGCGCTGGCTGCGCGCGGCGGACTCGGGAGGCCG CTCGGACGTACCTACTCGCTGCTGTGCTGCGCTGCGGCGCGAGGGCCCGCGGGG GCCTGCGAGGGGCGCGCGTGGCCTTCTACCGCGCCAGGCAGGGCTGCGGGAGCGAG CCGCCACGCTGCTGCACTGCGGCGCGCGCGCTACACCGTGGCGCTGGCGCGCGCT CAACGGCGTCTCGGGCCCGCGCGCGCGCGGGAACCACTACGCGCAGGTACCGCTC TCCACGGGCGCGGGGTAAGGCCGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCTG CGGCCCTGCGCCTCTTGGGGCGCGCCCGTGGTCTGCGGGATCAGCGCCCTGGGA GGAGGATGAGATCCGACGGGACCGAGTGAACCCAGAGCGTGTCCCTGCTGCGGGG GAGCCCATCCCTGCGGAGCCCTGGGGCCAATGACACGGAGTAGAGATCCGATACT ACGAGAAGGTGAGTGCGCAGAGTGAGCAGACTTACTCCATGGTGAAGACAGGGCGCC CACAGTCACCGTGATTTTCTCCAGCTGCTCAGGGTCCAGGGACACAGGCCCGCC ATTGTCGTACCGTAGTGACCATCTCGGCCCTCTCGTCTGGGCTCCGTGATGAGTG TGCTGGCCATTTGGAGGAGGAGGCCCTGCAGCTATGGCAAAGGAGGAGGGGATCCCCA TGATGAAGAGGAGCTGTATTTTCACTGTGAGTTGGCTGGGAAAGTCCCAACAGCTGCG ACATTCTGGACCCAGAGCTGTGGGACCTGCTGAGGCTGTGCATCTGTTCGCCA AGGAACTGGATGCCAAAGCGTCACGCTGGAGAGGAGCCTTGGAGGAGGCAAGCTGGG CGGGCGGTTTGGGGAGCTGTGCTGTGGCTGCTTGCAGCTCCCCGGTGCAGGAGCTG CTCGTAGCCGTGCATGCTGAGGGACAGCGCTCCGACTCACAGAGGCTCGGCTTCC TGGCCGAGGCCCTCAGCTGGGCCAGTTTGACCATAGCCACATCGTGGGCTGGAGGG CGTTGTTACCCGAGGTAGGGGAAGCACCTTGATGATTGTACCCGAGTACATGAGCCAT GGGGCCCTGGACGGCTTCTCAGGCAGCGGCACGAGGGGAGCTGGTGGCTGGGCAAC TGATGGGTTGCTGCTGGGCTGGCATCAGCCATGAAGTATCTGTGAGAGATGGGCTA CGTTACCGGGGCTGGCAGCTCGCCATGTGTGGTCAAGCAGCACTTGTCTGCAAG ATCTCTGGCTTGGGCGGGGCCCCGGGACCGATCAGAGGCTGTCTACACCACTATGG TGAGGCTACAGAGTGGCCGGAGCCAGCGCTATGGGCGGCTCCCGAGACCTTCAGTT TGGCCACTTCAGCTCTGCCAGTGACGTGTGGAGCTTCCGCATCATCATGTGGGAGGTG ATGGCTTTTGGGGAGCGGCTTACTGGGACATGTCTGGCCAGAGCTGGTGATCAAGG CTGTGGAGGATGGCTTCCGGCTGCCACCCCCAGGAAGTGTCTTAACCTTCTGCACCG ACTAATGCTCGACTGCTGGCAGAAAGGACCCAGGTGAGCGGCGCCAGGTTCTCCAGATC CACAGCATCTGAGCAAGATGGTGCAGGACCCAGAGCCCCCAAGTGTGCCCTGACTA CCTGTCCAGGCCCTCTGACCGCAGGCCCTCCCACTCCACTAGCCGACCGTGCCTTCTC CACCTTCCCTCTTTGGCTCTGTGGGCGCGTGGCTGGAGGCCCTGGACCTGTGCCG TACAAGGACAGCTTCCGCGCTGCTGGCTATGGGAGCCTGGAGGCCGTGGCCGAGATGA CTAGCCAGGACCTGGTGAGCCTAGGCATCTCTTGGCTGAACATCAGAGGCCCTCTCT CAGCGGGATCAGCGCCCTGCAGGCAGAGTGCTCCAGCTGCAGGGCCAGGGGGTGCG GTGTGAGTGGAACCCATTCTTCAAGGCAGGACTCCGGTGGGG </p>	

	ORF Start: ATG at 21	ORF Stop: TGA at 3078	
	SEQ ID NO: 112	1019 aa	MW at 110412.5kD
NOV34a, CG95175-01 Protein Sequence	MGACHHAWLIFVFLVEMGFHHVGQACLPLSLVILLDSKASQAE LGWTALPSNGWEEIS GVDEHDRPIRTYQVCNVLEPNQDNWLQTGWISRGGRQIFVELQFTLRDCSSIPGAAG TCKETFNVYYLETEADLGRGRPRLGKIDTIAADESFTQGD LGERKMKLNTEVREIGP LSRRGFHLAFQDVGACVALSVRVYYKQCRATVRGLATFPATAAESAFSTLVEVAGTC VAHSEGEPSPPRMHCGADGEWLVPVGRCSAGFQERGDFCEGICPPGFYKVSPPRP LCSPCEHSRALENASTFCVCQDSYARSPTDPPSASCTRPSPAPRDLQYSLRSPLVL RLRWLPPADSGGRSDVTYSLLCRCGREGPAGACEGPRVAF LPRQAGLRERAATLLHL RPGARYTVRVAALNGVSGPAAAAGTTYAQVTVSTGPGGKAVRAPHPEATAAPAPSW GRPVGPAGSAPWEEDEIRDRVEPQSVLSWREPI PAGAPGANDTEYEIRYKVSQAQ SEQTYSMVKTGAPTVTVIFLPAASGSRDQSPAIVTVVTISALLVLGSMVSLAIWRR RPCSYGKGGDAHDEELYFHCELAGKVPTRTFLDPQSCGDL LQAVHLFAKELDAKS VTLERSLGGGKLGRFGELCCGCLQLPGRQELLVAVHMLRDSASDSQLGFLAEALTL GQFDHSHIVRLEGVVTGRGSTLMIVTEYMSHGALDGF LRQRHEGQLVAGQLMGLLP LASAMKYLSEMGYVHRGLAARHVLVSSDLVCKISGFGRGPRDRSEAVYTTMVRLQSGR SPALWAAPETLQFGHFSSASDVWSFGIIMWEVMAFGERPYWDMSQGDVVIKAVEDGFR LPPPRNCPNLLHRLMLDCWQKDPGERPRFSQIHSILSKMVQDPEPPKCALTTCPRPLT RRPPTPLADRAFSTFPSPGSGVGAWLEALDLCRYKDSFAAAGYGSLEAVAEMTSQDLVS LGISLAHREALSGISALQARVLQLQGQGVQV		

Further analysis of the NOV34a protein yielded the following properties shown in Table 34B.

Table 34B. Protein Sequence Properties NOV34a	
PSort analysis:	0.4600 probability located in plasma membrane; 0.1000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 43 and 44

A search of the NOV34a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 34C.

Table 34C. Geneseq Results for NOV34a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV34a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM47209	Human NOV3 protein - Homo sapiens, 1000 aa. [WO200174851-A2, 11-OCT-2001]	7..1019 9..1000	933/1054 (88%) 936/1054 (88%)	0.0
AAU03553	Human protein kinase #53 - Homo sapiens, 1009 aa. [WO200138503-A2, 31-MAY-2001]	24..1019 18..1009	913/1048 (87%) 916/1048 (87%)	0.0
AAW03421	Mouse developmental kinase I - Mus sp, 998 aa. [WO9621013-A1, 11-JUL-1996]	7..1019 9..998	522/1055 (49%) 694/1055 (65%)	0.0
AAR85092	EPH-like receptor protein tyrosine kinase HEK11 - Homo sapiens, 998 aa. [WO9528484-A1, 26-OCT-1995]	7..1019 9..998	520/1055 (49%) 696/1055 (65%)	0.0
AAR85090	EPH-like receptor protein tyrosine kinase HEK7 - Homo sapiens, 991 aa. [WO9528484-A1, 26-OCT-1995]	28..1010 33..985	457/1018 (44%) 641/1018 (62%)	0.0

In a BLAST search of public sequence databases, the NOV34a protein was found to have homology to the proteins shown in the BLASTP data in Table 34D.

Table 34D. Public BLASTP Results for NOV34a				
Protein Accession Number	Protein/Organism/Length	NOV34a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAD13085	SEQUENCE 5 FROM PATENT WO0174851 - Homo sapiens (Human), 1000 aa.	7..1019 9..1000	933/1054 (88%) 936/1054 (88%)	0.0
CAD23752	SEQUENCE 1 FROM PATENT WO0208253 - Homo sapiens (Human), 974 aa (fragment).	33..1019 1..974	925/1027 (90%) 926/1027 (90%)	0.0
CAD23753	SEQUENCE 3 FROM PATENT WO0208253 - Homo sapiens (Human), 1022 aa (fragment).	33..1019 1..1022	930/1071 (86%) 931/1071 (86%)	0.0
CAD23754	SEQUENCE 5 FROM PATENT WO0208253 - Homo sapiens (Human), 647 aa (fragment).	33..653 1..623	575/661 (86%) 576/661 (86%)	0.0
CAD23755	SEQUENCE 7 FROM PATENT WO0208253 - Homo sapiens (Human), 695 aa (fragment).	33..653 1..671	580/705 (82%) 581/705 (82%)	0.0

PFam analysis predicts that the NOV34a protein contains the domains shown in the Table 34E.

Table 34E. Domain Analysis of NOV34a			
Pfam Domain	NOV34a Match Region	Identities/ Similarities for the Matched Region	Expect Value
EPH_lbd	32..203	115/178 (65%) 156/178 (88%)	4.4e-119
fn3	331..426	25/98 (26%) 66/98 (67%)	2e-06
pkinese	671..903	70/269 (26%) 167/269 (62%)	1.1e-37
SAM	942..1006	27/68 (40%) 50/68 (74%)	2.9e-15

Example 35.

The NOV35 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 35A.

Table 35A. NOV35 Sequence Analysis			
	SEQ ID NO: 113	765 bp	
NOV35a, CG95693-01 DNA Sequence	ATGGTGCTCTTTGGCTGGCTTTTTTCGCAGAGTTTCGCAGAGAGAATGTTATCTTTGGCA TTCTAACAGTGATGTCAATACAAGGTTATGCAAACCTCTGTAATCAATGGAGCATAAAC AGGAGAATTTAATGATTTGCCTCAGGAAGAACTTTACAGTGGATCAAATACAATACC GTACCAGGTGACAACCTCAGTGGGTCCAGCTCTTGCTGGACAGGCAGAGTGGTCTCTCC TCATTTCCACCACTACCACAGGCCCGTGGGAGGTAATGCCTGGCTACCGCCAACATTC ACTTAAGGCTGAAGGACCTTTAGTCAGCTTGTGGTGAGTGCTGCCAGGCCTGGAAC CACCTTCAGGGCAGTGGGCTCTCTCCAGCCCAGGGCAGATGCTGTCTTTGCAGGTG CCATGCCTACAATGGCAAGTGTCAAGCTGTCTACACTTCATCCCATTTGTGAATCATCC ACATTACGAAGATGCAGACTTGAGAAATTGTGAATTACGTTACTCACATGGAAAAAGG AAATTTGCAAATGTGATTAAATTAATGCTCATGAGATGTGGAGATTATCCTTATTATC CAGGGATCATGGGATGTGGTCTGCACACTCACTGAGAAGGTCCAGTAATGGCAAGGC AGTCACCAGGGGACCCATGTCTCCTCAGAGGACCTGGCTCCTCAGAGAAACCGCAGGTTT ACTGGGGGAGATTATCAGGACTTTAGACCTCAGGATGAAGAGACACTCCAGGAGGG GCTGTGGATAG		
	ORF Start: ATG at 1	ORF Stop: TAG at 763	
	SEQ ID NO: 114	254 aa	MW at 28531.2kD
NOV35a, CG95693-01 Protein Sequence	MVLFGWLFRRVRRENVIFGILTVMSIQGYANLCNQWSITGEFNDLPQEELLQWIKYNT VPGDNSVGPALAGQAEWSLLISTTTTGPEVMPGYRQHSCLKAEGPFSQLVVSAAAPGT HPSGQWALFQPRADAVFAGAMPTMASVKLSTLHPINVNHPHYEDADLRNCELRYSHGKR KFANVIKLMLMRCGDYPYPGIMGCLHTQLRRSSNGKAVTRGPMSSDLAPQRNRRF TGGDYQDFRPQDEETHSRRGCG		

Further analysis of the NOV35a protein yielded the following properties shown in
 5 Table 35B.

Table 35B. Protein Sequence Properties NOV35a	
PSort analysis:	0.5500 probability located in endoplasmic reticulum (membrane); 0.3592 probability located in lysosome (lumen); 0.2463 probability located in microbody (peroxisome); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 14 and 15

A search of the NOV35a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 35C.

Table 35C. Geneseq Results for NOV35a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV35a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAW85723	Novel protein (Clone AX56_28) - Homo sapiens, 171 aa. [WO9920644-A1, 29-APR-1999]	3..100 23..115	59/98 (60%) 64/98 (65%)	4e-23
AAG03191	Human secreted protein, SEQ ID NO: 7272 - Homo sapiens, 102 aa. [EP1033401-A2, 06-SEP-2000]	3..58 23..78	46/56 (82%) 48/56 (85%)	1e-19
AAM41346	Human polypeptide SEQ ID NO 6277 - Homo sapiens, 618 aa. [WO200153312-A1, 26-JUL-2001]	3..60 434..491	36/58 (62%) 45/58 (77%)	4e-15
AAM39560	Human polypeptide SEQ ID NO 2705 - Homo sapiens, 618 aa. [WO200153312-A1, 26-JUL-2001]	3..60 434..491	36/58 (62%) 45/58 (77%)	4e-15
AAG68230	Cycline-dependent human kinase 14 protein SEQ ID NO:2 - Homo sapiens, 129 aa. [WO200175013-A2, 11-OCT-2001]	130..180 5..56	38/52 (73%) 41/52 (78%)	3e-12

In a BLAST search of public sequence databases, the NOV35a protein was found to have homology to the proteins shown in the BLASTP data in Table 35D.

Table 35D. Public BLASTP Results for NOV35a				
Protein Accession Number	Protein/Organism/Length	NOV35a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96LZ9	CDNA FLJ32949 FIS, CLONE TESTI2008020, WEAKLY SIMILAR TO DPY-19 PROTEIN - Homo sapiens (Human), 758 aa.	3..100 576..668	59/98 (60%) 64/98 (65%)	1e-22
O94954	KIAA0877 PROTEIN - Homo sapiens (Human), 580 aa (fragment).	3..60 396..453	36/58 (62%) 45/58 (77%)	1e-14
S44629	F22B7.10 protein - Caenorhabditis elegans, 628 aa.	90..183 464..551	41/96 (42%) 51/96 (52%)	1e-09
P34413	Protein dpy-19 - Caenorhabditis elegans, 683 aa.	90..183 519..606	41/96 (42%) 51/96 (52%)	1e-09
Q9VWR8	CG6659 PROTEIN - Drosophila melanogaster (Fruit fly), 872 aa.	121..158 721..758	17/38 (44%) 22/38 (57%)	0.077

PFam analysis predicts that the NOV35a protein contains the domains shown in the Table 35E.

Table 35E. Domain Analysis of NOV35a			
Pfam Domain	NOV35a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 36.

- 5 The NOV36 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 36A.

Table 36A. NOV36 Sequence Analysis

	SEQ ID NO: 115	3500 bp
NOV36a, CG95814-01 DNA Sequence	GCGCTGCAGTTCTCCCGCTATGCATGGGCTGCGGCTCCGGCTCTAGCACAGGCACCA GCCGCGCCGCGCACCCGGCCCCAGCGCCACCGTCTGTCATGTGCCCGCCGTAGCCGTCT GCCAGCCCGCAGCCCGCGCTCCACGGAGCGCTGGAGACCACCGTGGGGGGCCCCCTC TGCCCTCGAGAGAAGCGGTCTTGGAGGTATGGATTTAGGTGGTTGGATTTTTCCTG GATCTATCAATTCAATTCGAATTTGGAAGAAAGAAAGAAACATGACGTCTCCAGC CAAAATCAAAAAGGATAAGGAGATCATAGCAGAGTACGATACTCAGGTCAAAGAGATC CGTGCTCAGTCCACAGAGCAGATGAAATGCCTGGACCAGCAGTGTGAGCTTCGGGTGC AACTGTTGCAGGACCTCCAGGACTTCTTCCGAAAGAGGCAGAGATTGAGATGGACTA CTCCCGCAACCTGGAGAAGCTGGCAGAACGCTTCCTGGCCAAAGACAGCAGCACAAG GACCAGCAATTCAAGAAGGATCAGAATGTTCTCTCTCCAGTCAACTGCTGGAATCTCC TCTTAAACCAGGTGAAGCGGGAAGCAGGGACCATAACCCTTGAGTGACATCTACCT GAATAATATCATTCTCGATTTGTACAAGTCAGCGAGGACTCAGGAAGACTCTTTAAA AAGAGTAAAGAAGTCGGCCAGCAGCTCCAAGATGATTGATGAAGGTCTCGAACCAGC TCTACTCGGTCTAGAACATATCAGTGTACAATGCCACAGCATCAGTGTCTCAGAG CAAACTAAAGGAGCGGAGAAGCAGGAGGAGAAGCAAATTTGGTAAATCGTAAGCAG GAGGACCGGCAGACCCACGCTCCCTGACTCCACGCCAACGTTCCGATTGAGGAGA AACATGTCCGGAGGAGTCTAGTGAAGAAGATTGAGAAGATGAAGGAGAAGCACAAGC CAAGTACACGGAGATAAGCTGAAGGCCATCAAAGCCCGAATGAGTACTTGTCTGGCT TTGGAGGCAACCAATGCATCTGTCTTCAAGTACTACATCCATGACCTATCTGACCTTA TTGATCAGTGTGTGACTTAGGCTACCATGCAAGTCTGAACCGGGCTCTACGCACCTT CCTCTCTGCTGAGTTAAACCTGGAACTGCAAGCATGAGGGTCTGGATGCCATCGAG AATGCAGTAGAAAACCTGGATGCCACCATGACAAGCAGCGCCTCATGGAGATGTACA ACAACGTCTTCTGCCCCCTATGAAGTTTGAAGTTTCAGCCCCACATGGGGGATATGGC TTCCAGCTCTGTGCCAGCAGCTGTCCAGAGTGAGCTGGTACAGAGATGCCAACA CTGCAGTCTCGCTTATCCACTCTAAAGATTGAAAACGAAGAGGTAAAGAAGCAATGG AGGCCACCTCGCAACCATCCAGGACATGTGACTGTGAGGACTTTGATGTGTCTGA CTGCTTCCAGTACAGCAACTCCATGGAGTCCGTCAAGTCCACGGTCTCTGAAACCTTC ATGGGCAAGCCAGCATTGCTAAGAGGAGAGCAACAGCAAGAGACAGAGCAGTTTT ATTTCACAAAAATGAAGAGTACCTGGAGGGCAGGAACCTCATCAAGTTTACAAGC CAAGCATGACCTTCTGCAGAAAACCTGGGAGAAAGTCAGCGGACAGATTGCACTTA GCCAGGCGCAGCTCAACTGTGAGGAAACAGGACTCCAGCCAGGCAATCTCTAGTGG TGGAAAGCTGTATCCGTTTATCAGCAGACCGGACTACAGCATGAAGGAATTTCCG GGTGTGAGGATCCAGGTGGAAGTGAATGACATCAAAAATGCTTTGAGAGAGGAGAG GACCCCTGCTGGGGACAGAACGACCATGACATGGATTCCATAGCTGGTGTCTCTGA AGCTTTACTTCCGGGGGCTGGAACACCTCTCTTCCCAAGGACATCTTTATGACCT GATGGCTGCGTCAAAATGGACAACCTGCAGGAGAGAGCTCTGCATCCGGAAGATC CTCTAGTCTGCCCCAAACCACTCTGATTATCATGAGATACCTCTTGGCTTCTCTCA ATCATTTATCAGTTTCAAGTGAAGAGAACATGATGGACCCCTACAACTCGCCATCTG CTTCGGGCCCTCGCTAATGTCAAGTCCAGAGGGCCACGACAGGTGTCTGCAAGCC CACGTGAATGAGCTGATCAAAACCATCATCATCCAGCATGAGAACATCTTCCCAAGCC CCAGGGAGCTGGAGGGCCCTGTCTACAGCAGAGGAGGAAGCATGGAGGATTACTGTGA TAGCCCTCATGGAGAGACTACCTCGGTTGAAGACTCAACCCAGGATGTGACCCGAGAG CACCACAGAGCGATGACGAATGTGAGCCCATCGAGGCCATTGCCAAGTTTGACTACG TGGGCGGACAGCCCGAGAGCTGTCTTTAAGAAGGGAGCATCCCTGCTGCTTTACCA GCGGGCTTCCGACGACTGGTGGGAAGGCCGGCACAATGGCATCGACGGACTCATCCCC CATCAGTACATCGTGGTCCAAGACACCGAGGACGGTGTGCTGGAGAGGTCCAGCCCCA AGTCTGAGATTGAGGTCAATTTCTGAGCCACCTGAAGAAAAGGTGACAGCCAGAGCGGG GGCCAGCTGTCCAGTGGGGGTCTAGTACCGATATTTATCTTGCAACATCAACAAG CAAAGGAAGCGTCCAGAATCTGGGAGCATCCGGAAGCTTTTCCGAGTGACAGCCATG GGCTGAGCAGTTCCTGACTGACTCTCTCTCCAGGGGTGGGGGTAGTGTCCGCC ATCTCCAGCCCATCATGAGCCAGAGCTCCCCAAGAGAGGGCCAGATAAGTGTTC ATCAGTGGGACGGGAGCCTCAACTCCATCAGCCGCACTCATCCCTGAAGAATCGGC TGGATAGTCCACAGATCCGGAAGACTGCCACAGCGGGAAGGTCAAAAAGCTTCAATAA CCATCGGCCATGGACCTGAGGTCAATGCTCAGGATATTGAGGCAACATGAATCG GCCCTGAATGAGCTACGGGAAGTAGAACGGCAGAGCAGTGTCAAACACACCCCTGACG TGGTTCTGGACACCTTGGAGCCCTCAAACCTCCCAAGTGGTGGCCCCCAGCTCAGA GCCCTCCAGCCCTCTGCACACCCAGCTCTCAAGGACCCCGAGCCCGCTTCCAGCGC AGCGCCAGTACTGCTGGGACATCGCTGCGCTTCCGGCCTGTCAAGTGTCAAGA TGGCTGCCCGGTCAAACCAACGACCCACAGGCCCAAGCCCACTGTCTTCCCAAAAC AAATGCCACTAGCCCTGGTGTCAACTCATCAACTTCCCCACAGTCTACTGACAAGTCT TGTAAGTCTGAGGGATAAT	

	ORF Start: ATG at 25	ORF Stop: TGA at 3490	
	SEQ ID NO: 116	1155 aa	MW at 130304.7kD
NOV36a, CG95814-01 Protein Sequence	MCGCGSGSSTGTSTRRRTRPQRPPSACARRSRLPSPQPALHGALETTVGGPFCPREKRSW RYGFRWLDFFRGSINSQFEFGRKKENMTSPAKFKDKKEIIAEYDTQVKEIRAQLTEQM KCLDQQCELRVQLQLQDLQDFFRKKAIEIMDYSRNLEKLAERFLAKTRSTKDQQFKKQDQ NVLSPVNCWNLLLNQVKRESRDHTLSDIYLNIIIPRFVQVSEDSGRLFKKSKEVGQQ LQDDLKVLNELYSMKTYHMYNADSSISAQSKLKEAEKQEEKQIGKSVKQEDRQTPRS PDSTANVRIEEKHVRRSSVKKIEKMKEKHQAKYTENKLKAIKARNEYLLALEATNASV FKYYIHDLSDLIDQCCDLGYHASLNRALRTFLSAELNLEQSKHEGLDAIENAVENLDA TSDKQRLMEMYNNVFCPPMKFEFQPHMGDMASQLCAQQPVQSELVQRCCQLQSRSLSTL KIENEEVKKTMEATLQTIQDIVTVEDFDVSDCFQYSNMSVSKSTVSETFMGKPSIAK RRANQQETEQFYFTKMKEYLEGRNLITKLQAKHDLQKTLGBSQRTDCSLARRSSTVR KQDSSQAIPLVVESEIRFISRHGLQHEGIFRVSGSQVEVNDIKNAFERGEDPLAGDQN DHDMDSIAGVLKLYFRGLEHPLFPKDI FHDLMACVTMDNLQERALHIRKVLVLPKTT LIIMRYLFAFLNHLNQFSEENMDPYNLAI CFGPSLMSVPEGHDQVSCQAHVNELIKT IIIQHENIFPSPRELEGPVYSRGGSMEDYCDSPHGETTSVEDSTQDVTAEHHTSDDEC EPIEAIKFDYVGRRTARELSFKKGASLLLYQRASDDWEGRHNGIDGLIPHQYIVVQD TEDGVVERSSPKSEIEVISEPPEEKVTARAGASCPSGGHVADIYLANINKQRKRPESG SIRKTRFRSDSHGLSSSLTDSSSPGVGASCRPSSQPIMSQSLPKEGPDKCSISGHGSLN SISRHSLLKNRLDSPQIRKTATAGRSKSFNNHREMPDPEVIAQDIEATMNSALNELREL ERQSSVKHTPDVVDLTLEPLKTSPPVVAPTSEPSSPLHTQLLKDPEPAFORSASTAGDI ACAFRPVKSVMKMAAPVKPPATRPKPTVFPKTNATSPGVNSSTSPQSTDKSCTV		

Further analysis of the NOV36a protein yielded the following properties shown in Table 36B.

Table 36B. Protein Sequence Properties NOV36a	
PSort analysis:	0.9400 probability located in nucleus; 0.4936 probability located in mitochondrial matrix space; 0.3000 probability located in microbody (peroxisome); 0.2087 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV36a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 36C.

Table 36C. Geneseq Results for NOV36a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV36a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM79887	Human protein SEQ ID NO 3533 - Homo sapiens, 1094 aa. [WO200157190-A2, 09-AUG-2001]	62..1155 1..1094	1088/1094 (99%) 1090/1094 (99%)	0.0
ABG18053	Novel human diagnostic protein #18044 - Homo sapiens, 1095 aa. [WO200175067-A2, 11-OCT-2001]	63..1155 3..1095	1089/1093 (99%) 1089/1093 (99%)	0.0
ABG05128	Novel human diagnostic protein #5119 - Homo sapiens, 1095 aa. [WO200175067-A2, 11-OCT-2001]	63..1155 3..1095	1089/1093 (99%) 1089/1093 (99%)	0.0
ABG18053	Novel human diagnostic protein #18044 - Homo sapiens, 1095 aa. [WO200175067-A2, 11-OCT-2001]	63..1155 3..1095	1089/1093 (99%) 1089/1093 (99%)	0.0
ABG05128	Novel human diagnostic protein #5119 - Homo sapiens, 1095 aa. [WO200175067-A2, 11-OCT-2001]	63..1155 3..1095	1089/1093 (99%) 1089/1093 (99%)	0.0

In a BLAST search of public sequence databases, the NOV36a protein was found to have homology to the proteins shown in the BLASTP data in Table 36D.

Table 36D. Public BLASTP Results for NOV36a				
Protein Accession Number	Protein/Organism/Length	NOV36a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
O75044	KIAA0456 PROTEIN - Homo sapiens (Human), 1095 aa (fragment).	63..1155 3..1095	1089/1093 (99%) 1089/1093 (99%)	0.0
Q9P2P2	KIAA1304 PROTEIN - Homo sapiens (Human), 1051 aa (fragment).	107..1155 12..1051	719/1063 (67%) 850/1063 (79%)	0.0
CAC22407	SEQUENCE 47 FROM PATENT WO0075320 - Homo sapiens (Human), 1075 aa.	85..1151 1..1062	642/1079 (59%) 819/1079 (75%)	0.0
Q92512	RHO GTPASE-ACTIVATING PROTEIN - Mus musculus (Mouse), 987 aa.	204..1151 8..974	554/983 (56%) 713/983 (72%)	0.0
Q91Z69	GAP1 - Mus musculus (Mouse), 714 aa (fragment).	447..1155 3..714	439/721 (60%) 545/721 (74%)	0.0

PFam analysis predicts that the NOV36a protein contains the domains shown in the Table 36E.

Table 36E. Domain Analysis of NOV36a			
Pfam Domain	NOV36a Match Region	Identities/ Similarities for the Matched Region	Expect Value
FCH	106..204	27/109 (25%) 81/109 (74%)	2.2e-19
RhoGAP	589..741	54/170 (32%) 128/170 (75%)	4.4e-53
SH3	815..869	20/58 (34%) 41/58 (71%)	1.7e-11

Example 37.

- 5 The NOV37 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 37A.

Table 37A. NOV37 Sequence Analysis

	SEQ ID NO: 117	3790 bp
NOV37a, CG95824-01 DNA Sequence	CAAACAGCGATGTCCAGACGCGAGGACTACGAGTGCAGGAGCCATAATGTGCACTGCG CGGAGTCGAGGATTCCAGGGTCGAACACTCGGTTGGAGTGGGTGGAGATCATCGAACCC GCGCACCCGCGAGCGCATGTACGCCAACCTGGTCACCGGTGAGTGCCTGTGGGACCCG CCGGCCGCGTCCGCATCAAGCGCACGAGCGAGAACCAGTGGTGGGAGCTGTTCCGACC CCAACAGCTCCCGCTTCTACTACTACAATGCCAGCACGCGACGCGTGTGGCACCG GCGCGAGGGCTGCGACATCATCCCGCTGGCCAAGCTGCAGACGCTGAAGCAGAACACG GAGTCCCGCGCGCTCGGCGGAGAGCAGCCCCGGGCGCGGAGCAGCGTCAAGCGCTG AGGGCAGCACCAGCTCTCTCTGGAGCCCGAGCCGACACTGAGAAAGCGCAGAGTT GCCAGCGAGGGCCGGGCGGCGCGGCGGCTTGGGACAGTGAAGGAGGACAGCGGAGC TCTTCAACCACAGGAGTGTCTCTTGAAGAAGGACTATGAGATTACCGGGATTACAGTG CGGACGGCCAGCTTCTTCACTACAGGACCTCTCTCGCTGCGGTGGAACTCGGGCGCCAA AGAGCGCATGCTCATCAAGGTGCTGATCGGGAGCCAGCTTCTCGCGCGCCAGGGC AATGGCTACGCCCCAGACGCCCCACTGGGGTCCGCTCCCGCAGACCTCCGGCAGCC AGCACTCACCAGCCTGCAGACCTTCGCCCCGAGGGCTGACGGCACCATTCTTCTCC AGAGAGGAGGCGCTCACCTTCTGAAGAGGGCCGAGCTCCAGGCGGAGCTCCCGC CTGCTGGCCAGCCCCGAAAGCCCTCGGGGACTCGCAGCCCTCTCCCGCGCTATG GCTATGAACCCCGCTTCTACGAGGAGCCCCAGTGGAGTACAGGCCCCCATCTACAGA TGAGCCCCCATGGACGTGCAATTGAGGCTGGCGGGGCTACGAGCCGCTCTCC CAGCGGTGCGCGGGCGTAAGCCCCGGCGGCTCTCTCAGCCCCAACAGAGGGCCCCC CCTCGCCTGCGCAGCAGCTGGTGTCTACCAAGCAGAAGTGTCCCGAGCGCTTCTGAG CCTGGAGTACAGTCCCGCCGGAAGGAGTACGTGCGGCGAGTGGTCTACGTGGAGCAG GCGGGCTCCAGCCCCAAGCTGCGCGCGCGCCCGCGGCACAAGTACGCGCCCAACCCG GCGGTGGTTCGTACTCTTGCAGCCAGCCCTGCCTGCTGAGGGACCAGCGCTGGG CGTCAAGTCCGAGACTACAGCACCATTGAGGGGACCTGAGCTCGGCGACAGCCAGCCG CCCACGCGCTGCCACAGGCCAGGAGGATGCCATGCTCTGGTCCAGCCAGCAGGACA CCTGTCTCTCACAGGCTACTCCCCGGGCAGCGCAAGCGGAAGAGCAGAAAGCCCTC TTTGTGCCAAGCCACCAGCGCCACCCCACTGAGGGCCCCGGGGACCTGTCTGTGGAG CAGCCCCCTGGCCGAGGAACAGCCCCGCTGCGGGACAGCCTCGCCCCGCTGAAGCGAG CGGAAGGTGAGGCCGAAGGGCGCGGGGCGCGGCGAGCCCTTCTGGCGCAGGCTCG GCTGGCCTGGGAGGCGCAGCAGGCCCACTTCCACATGAAGCAGAGGAGCAGCTGGGAC TCCAGCAGGACGGCTCTGGCTACGAGAGCGAGCGCGCCTGCCATGCCCATGCCCG GGCGGTGGTGCGGGCTTCAGCGAGGACGAGGCGCTGGCCAGCAGGAGAACAGGCA CTGGAGGAGGGGCACTTCGAGAAGCTAGGCTTCCCCAGATCTGCTGGAGAGAGC GTCTCGTGACAGCAACCTGGCCTCACCAGAGCCCTACTCCACCCCTCAGAGTGTG AGGACCTCGCTGCTGTGCCAGTTTCGAGAGCAGCCGGCAGAGCCCGCAGCGCGCTTCC CAGCTCCAGCTGCTTCTCCCACTTTCAGCTGCGCAAGCCCTCTCTCGGAGCGGAC ATCGAGAAGTGGGCTTCCAAGCACTTCAACAAGCACACGAGGGCCTCTTCCGGCGGA AGGTGTCCATCGCCAACATGCTGGCTGGAGCAGCGAGTCCATCAAGAAGCCCATGAT CGTGACAAGCGACCGGCACGTGAAGAAGGAGGCTGCGAGCTCTTCAAGCTGATCCAG ATGTACATGGGTGACCGCGGGCCAGGGCCGACCCACTGCACGTGCTGGAGGTGG CCACCAAGGGCTGGAGCGTGCAGGGCCTGCGGGACGAGCTCTACATCCAGCTGTGCCG GCAGACCACCGAGAAGTTCGCGCTGGAGAGCCTGGCCCGCGGTGGGAGCTCATGGCC ATCTGCTGGCCTTTTCCCGCCACCCCAAGTTCCTACTCTACCTGGAAGGCTACA TCTACCGGCACATGGACCCCGTCAATGACACTAAAGGGGTGGCGCTAAGCAGTATGC CAAGTACTGTACCACAAGCTACAGAAGGCAGCCCTGACCGGGCCAGAAGGGGCTG AAGAAGCCCAACGTGGAGGAGATCCGGCATGCCAAGAAGCCCGTGTTCAGCCCGTCCA TGTTCGGCAGCGCACTGCAGGAGGTATGGGCATGCAGAGAGAGCGCTACCCCGAGCG CCAGCTGCCCTGGGTGCAGACACGGCTCTCTGAGGAGGTGCTGGCGCTCAACGGTGAC CAGACAGAGGGCATCTTCAGGGTCCCTGGGGACATTGACGAGGTGAATGCCCTGAAGC TGCAAGTGGACAGTGAAGGTGCCACAGGCCCTGGAAGACCCCAAGTCCCTGCGTCTC CCTGCTGAAGCTGTGGTACCGGGAGCTGGAGGAGCCCTGATCCCGCAGAGTTCTAC GAGCAGTGCATCGCGCACTACGACAGCCCCGAGGCGCGGTGGCGGTGGTGCACGGCG TGCCCCGATCAACCGCATGGTGTGTGTACCTCATCCGCTTCTTGAGGTCTTCTGT GCGCCCAACTGTCTGCGCTGCCAGTCCGACGACCGCGCGTCTCTCGAAGAACCCC GCAAGGAGATGTCTTCTGCGGGTGTCTATCCAGCACCTGGACACAGCTTCATGGA GGGTGTGCTGGAGCGGGGCGCCCGGGGACAGGAGGATGTCTGCGCGCCCCAGCCA GGCAGAACTCCGCACTCGCTCTCCCGCAGAGGGGCCAGAACTCCCGCGCCAGCCCT GGAGCCCCCTCACTCCCCAGGCCCTGGCCCCGGCGCTCCCCAGCTTCTTGTGCTG GTCTGAGGGTGCAGCCAGGGCAGCAGCGCGCGGGAGGGCGCTCTGGCCCCCACC TCAGGGCCAGTTCCCGGGCACCGCCTCGCCCTCGCTGGCGCGGGTCACTCCGA GAAAGTGCCTTCTGTCTGGAGCGAGCGAGCTGCTCTCTTGGGGCGGGGTGCTGCT TCCCTGTGGCTCTGCGCGCCCTGGCTGGGCTTGGCCAGCCCGCCCGGTCTCTCT TCCCTTTCTCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT CGACCTGTCCCGCT CAGGAGTTTGGGGGCTCCAG	

	ORF Start: ATG at 10	ORF Stop: TGA at 3763	
	SEQ ID NO: 118	1251 aa	MW at 138786.4kD
NOV37a, CG95824-01 Protein Sequence	MSQTQDYECRSHNVLDPESTRIPGSNTRLEWVEIIIEPRTRERMYANLVTGECVWDPPAG VRIKRTSENQWELFDPNSTRFYYNASTQRTVWHRPQGCIIPLAKLQTLKQNTESP RASAESSPGRGSSVSREGSTSSSLEPEPDTEKAQELPARAGRPAAGFTVKEDSGSSSP PGVFLEKDYEIYRDYSADGQLLHYRTSSLRWNSGAKERMLIKVADREPSFLAAGNGY APDGPPGVRSRPSGSHPSLQTFAPADGTIFFPERRPSPFLKRAELPGSSSPLLA QPRKPSGDSQPSSPRYGYEPPLYEPPVEYQAPIYDEPPMDVQFEAGGGYQAGSPQRS PGRKPRPFLQPNKQPPSPCQQLVLTQKQKPERFLSLEYSPAGKEYVRQLVYVEQAGS SPKLKAGPRHXYAPNPGGGSYSLQSPCLLRDQRLGVKSGDYSTMEGPELRHSQPPTP LPQAQEDAMSWSSQDQLSSTGYSPTGRKRKSRKPSLCQATSATPTGPGDLLVEQPL AEEQPPCGTSLAPVKRAEGEAGARGAAEPFLAQARLAWAQAQAHFHMQRSSWDSQQ DGSGYESDQALPLPMPGPVVRAFSEDEALAQQENRHWRRTFEKLGFPQILLEKSVSV QTNLASPEPYLHPSQSEDLAACAQFESSRQSRSGVPSSSCVFPTFTLRKPSSETDIEN WASKHFNKHTQGLFRRKVSIAANLAWSSSEIKKPMIVTSDRHVKKEACELFKLIQMYM GDRRAKADPLHVALEVATKGWSVQGLRDELYIQLCROTTFNFRLESLARGWELMAICL AFFPPTPKFHSYLEGYIYRHMDPVNDTKGVAISTYAKYCYHKLQKAALTGAKKGLKKP NVEEIRHAKNAVFSMFGSALQEVGMQREYPERQLPWVQTRLSEEVLALNGDQTE GIFRVPDIDEVNALKLQVDQWKVPTGLEDPHVPASLLKLWYRELEPLIPHEFYEQC IAHYDSPEAAVAVVHALPRINRMVLCYLIRFLQVFPQANNAVTKMDVSNLAMVMAPN CLRCQSDDPVIFENTRKEMSFRLVLIQHLDTSMFEGVLERGRPGTGGMSCRPQPGRT PHSLSRQRGQNRPAQPWSPLHSPRLAPALPTSSAWSEGAARAQRRGRLWPPTSRLP VPAGTASPSAGRGSAPRKCLLCPGAERRCLLGAGLPPCGSCAPWPGPCPAAPVSPSLP PVLVLACSSSQPRESFPTCPRLLSLPRPVPSW		

Further analysis of the NOV37a protein yielded the following properties shown in Table 37B.

Table 37B. Protein Sequence Properties NOV37a	
PSort analysis:	0.7000 probability located in nucleus; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV37a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 37C.

Table 37C. Geneseq Results for NOV37a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV37a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB59882	Drosophila melanogaster polypeptide SEQ ID NO 6438 - Drosophila melanogaster, 1309 aa. [WO200171042-A2, 27- SEP-2001]	681..1083 897..1309	213/418 (50%) 282/418 (66%)	e-114
AAE03048	Human preoptic regulatory factor-2 (hPORF-2) protein #1 - Homo sapiens, 75 aa. [WO200142464-A2, 14-JUN- 2001]	1009..1083 1..75	75/75 (100%) 75/75 (100%)	7e-36
ABG06669	Novel human diagnostic protein #6660 - Homo sapiens, 322 aa. [WO200175067-A2, 11-OCT-2001]	872..1091 28..253	76/230 (33%) 111/230 (48%)	2e-17
ABG06669	Novel human diagnostic protein #6660 - Homo sapiens, 322 aa. [WO200175067-A2, 11-OCT-2001]	872..1091 28..253	76/230 (33%) 111/230 (48%)	2e-17
AAB42926	Human ORFX ORF2690 polypeptide sequence SEQ ID NO:5380 - Homo sapiens, 903 aa. [WO200058473-A2, 05- OCT-2000]	928..1250 84..447	100/377 (26%) 142/377 (37%)	2e-13

In a BLAST search of public sequence databases, the NOV37a protein was found to have homology to the proteins shown in the BLASTP data in Table 37D.

Table 37D. Public BLASTP Results for NOV37a				
Protein Accession Number	Protein/Organism/Length	NOV37a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9C0H5	KIAA1688 PROTEIN - Homo sapiens (Human), 1094 aa (fragment).	1..1083 12..1094	1083/1083 (100%) 1083/1083 (100%)	0.0
Q9VDE9	CG3421 PROTEIN - Drosophila melanogaster (Fruit fly), 1309 aa.	681..1083 897..1309	213/418 (50%) 282/418 (66%)	e-113
P18890	Putative preoptic regulatory factor-2 precursor (PORF-2) - Rattus norvegicus (Rat), 75 aa.	1009..1083 1..75	74/75 (98%) 75/75 (99%)	3e-35
Q9BKW0	HYPOTHETICAL 36.4 KDA PROTEIN - Caenorhabditis elegans, 317 aa.	798..932 179..307	50/135 (37%) 80/135 (59%)	3e-20
AAH23344	SIMILAR TO HYPOTHETICAL PROTEIN DKFZP564B1162 - Mus musculus (Mouse), 654 aa.	872..1091 22..247	74/230 (32%) 109/230 (47%)	4e-17

PFam analysis predicts that the NOV37a protein contains the domains shown in the Table 37E.

Table 37E. Domain Analysis of NOV37a			
Pfam Domain	NOV37a Match Region	Identities/ Similarities for the Matched Region	Expect Value
WW	65..95	12/31 (39%) 19/31 (61%)	0.0054
MyTH4	761..879	34/123 (28%) 85/123 (69%)	1.4e-19
RhoGAP	909..1056	52/171 (30%) 105/171 (61%)	7.3e-19

Example 38.

- 5 The NOV38 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 38A.

Table 38A. NOV38 Sequence Analysis

	SEQ ID NO: 119	1269 bp	
NOV38a, CG96198-01 DNA Sequence	CTTTATGTTCCATTTAGGTGAAGAAAGCCCTGTGTGTCCTCGTCCAACACCGGTCTT ACAAAGCCCCCGGGCAGCATCCTTTTCTCGCCACACTGCAGCAAGCTCCTCCCTTA TTGAAGCAGAAGCTCCACCTTCTCGGCAGCAGCCAAGGTCGAGTGGGAGTGACCA TGTGCTGTTTGTTCCTTCGGCTGGAGGGAACCTGGCAGTCTGTGGTCAGTGACACG CGGGAACCCAGCCAGCTCCGCCCCCTCTCCGGTCAGGAGACTCTTCATCCATTAGCA CCGCCGACTTACTAGTCACTTTTGTGGTTATGCCCTAGATGCCACTGGAATATCA TGTTCAATGGCTGGCTGTGGACATCGCATGTGCGACACTGATGTTCTGAAACTAAT GCCACGTATTCTGCAGCTTTCCTGCCTGTGGTCATTGGATTGGACCGCCAGGCAGCA TACTCAACCCGCTTGGATCCCGTTCAAGGTGTAAAGGAACTTCTGGGGGCGAGCTGGG ACTTAGTTTCTGCTTGCCTTCCCCAGGTGGCTGGCCAGTCCCTTTCACTCAGTGT GTCAACCAAGGCAGCTTCAAGGCTCAATGGCAAGAGACCACCTATAACCTCTTCACCT TCTGCTGCCTCTTCTGCTGCCACTGACTGCCATGGCCATCTGCTATAGCCGCATTGT CCTCAGTGTGTCCAGGCCCCAGACAAGGAAGGCTGGTGAATTTGCCCTCCCCGCTCC TTTGACAATTGTCCCCGTGTTCTGCTCTCCGGGCCCTGAGACTGGCCCTGCTTATCTGT TGACCTTCATCCTCTGTCTGGACACCTTATTACCTACTGGGTATGTGGTACTGTTCTG CCCCACCATGCTAACTGAAGTCCCTCCAGCCTGAGCCACATCCTTTCTCTTGGGG CTCCTCAATGCTCCTTTGGATCCTCTCCTCTATGGGGCCTTCACCTCCGCCCGCCTCA GCCTCCAAAGTGTGGGATTACAGGCGTGAGCCACCGCGCCAGCCAGGCTTTGACTA TTTTTATACTCCTCACAATTCTCAAGCACTTCTTAGGCAGAGCACAAAAACAAAAAA AAAAAAAACCTAGAAATATATGAATGCAAGAGATGGGAAGAGGAAAGACCTTTCC AACTAGCTCAGAGTAGCCTGATCGTAACTCCAAACAGTTCAAAAATTTCTTTAGGCT GAAACTATTTCATCTTCTCTAAATGTCCCTCATGTCTGCTAAAATCACTTTC		
	ORF Start: ATG at 5	ORF Stop: TGA at 1124	
	SEQ ID NO: 120	373 aa	MW at 41692.8kD
NOV38a, CG96198-01 Protein Sequence	MFHLGEESPVCPRPTVLTKPPGQHPFLATLQQAPPLVEAEAPTFSAAKVRVGVTVI LFVSSAGGNLAVLWSVTRRPSQLRPSVRRLFILHAAADLLVTFVVMPLDATWNITV QWLAVDIACRTLMFLKLMATYSAAFLPVVIGLDRQAALNPLGSRSGVRKLLGAAWGL SFLLAFPQVAGVPVPTQCVTKGSFKAQWQETTYNLFTFCCLFLLPLTAMAIYSRIVL SVSRPQTRKAGEFALPRSFDCNPRVRLRALRLALLILLTFILCWTPYYLLGMWYWFSP TMLTEVPPSLSHILFLLGLLNAPLDPLLYGAFTSARLSLPKCWDYRREPPRPASFDYF YTPHNSQALLRQSTKTKKKKKLRNI		
	SEQ ID NO: 121	1551 bp	
NOV38b, CG96198-02 DNA Sequence	ATGTTTATCTCATCCATTACAAGGTAAGTTCAGAGGAGCAGAGGCCTGTTTGTCT TTTCTTTTTTTGGAGACGGTATTTCACTGTTGTGCCCCAGGCTAGAGTACAATGGCA CGATCTCGGCTCACCACAACTCTGCCTCCCGGGTCAAGCGATTCTCTGCCTCAAC CTCCCGAGTAGCTGGGATTACAGGCATGCGCCACCACACCCGGCTAATTTGTATT TAGTAGAGACGGGTTTCTCTGCCACCTTCTCGGCAGCAGCAAGTCCGAGTGGG AGTGACCATTTGTGCTGTTTGTTCCTTCGGCTGGAGGGAACCTGGCAGTCTGTGGTCA GTGACACGGCGGGGAACCCAGCCAGCTCCGCCCTCTCCGGTCAGGAGACTCTTCATCC ATTTAGCAGCCGCGACTTACTAGTCACTTTTGTGGTTATGCCCTAGATGCCACCTG GAATATCACTGTTCATGGCTGGCTGTGGACATCGCATGTGCGGACACTGATGTTCTCT AACTAATGGCCACGTATTCTGCAGCTTTCCTGCCTGTGGTCATTGGAATTGGACCGCC AGGCAGCAGTACTCAACCCGCTTGGATCCCGTTCAAGTAAAAGAAAAAGAAAGCAAT TATGGTTCTTGGCTGGGCGCTGAGGTTCAATGCCAGTATCCAGCACTTTGGAAGGCT GAGGTGAGCAGATCAGGCCAGTCCCTTTCACTCAGTGTGTACCAAAGGCAGCTTC AGGCTCAATGGCAAGAGACCACCTATAACCTCTTCACCTTCTGCTGCCTCTTCTGCT GCCACTGACTGCCATGGCCATCTGCTATAGCCGCAATTGTCTCAGTGTGTCCAGGGAA TTTGCCCTCCCCCGCTCCTTTGACAATTGTCCCGTGTTCGTCTCCGGGCCCTGAGAC TGGCCCTGCTTATCTTGTGACCTTCATCTCTGCTGGACACCTTATTACCTACTGGG TATGTGGTACTGGTTCTCCCCACCATGCTAACTGAAGTCCCTCCAGCCTGAGCCAC ATCCTTTTCTCTTGGGCTCCTCAATGCTCCTTTGGATCCTCTCTCTATGGGGCCT TCACCTTGGCTGCCAAGATTATTTCACTGGATCTCTGAATACCAAGCCCTCCAC CATGGCCAGCCGGGTGGGGGCCGGGTCGTGGCCGGGGCCAGTTGACCTTCAACGTG GAGGCCGTGGGCATTGGGCTCCACAGTAGGTACTGTGATCATTTAAAAATCAGTTGGG TTACATACCTACTAAAACTCTTGTGTTGTTGTTTGGAGCAGTAGTCTCGCTCTATC ACCCAGGCTAAGAGTCAGTGGTGCAGTCTCGGCTCACTGCATCTCTGCTCCCGGG TTCAAGCAATTCTCTGCCTCAAGCCTCTGAGTACCTGGGACTACAGGCGCCCTCCA AGCCTCTGAGTACCTGGGACTACAGGCGCCCCCCCCACACACACACACACACCCGG CTAATTTTGTATTTTAGTAGAGACGGGATTTACCATGTTG		

	ORF Start: ATG at 1	ORF Stop: TAA at 1510	
	SEQ ID NO: 122	503 aa	MW at 56199.8kD
NOV38b, CG96198-02 Protein Sequence	MFISSHYKVSSRGAEALFVFSFFLRRYFTVVAQARVQWHDLGSPQTLPPGFKRFSCLN LPSSWDYRHAPPHANFVFLVETGFLLPFSAAKVRVGVTIVLFVSSAGGNLAVLWS VTRREPSQLRPSVRRLLFIHLAADLLVTFVVMPLDATWNITVQWLAVDIACRTLMFL KLMATYSAFLPVVIGLDRQAAVLNPLGSRSGKRKKKAIMVPGWALRFMPVIPALWKA EVSRSQVPVPTQCVTKGSFKAQWQETTYNLTFCCLFLLPLTAMAI CYSRIVLSVSRE FALPRSFDCNCPVRLRALRLALLLTFILCWTPYYLLGMWYWFSPMTL TEVPPSLSH ILFLLGLLNAPLDPLLYGAFTLGCRRLFTGSLNTQAPSTMASRGGRGRGRGQLTFNV EAVGIGLHSRYCDHFNQLVTYLLKILVVVVLRRSLALSPRLRVQWCSLGS LHPLPPG FKQFSCLKPPEYLG LQAPSKPPEYLG LQAPPPTHHTPG		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 38B.

Table 38B. Comparison of NOV38a against NOV38b.		
Protein Sequence	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV38b	43..323	214/291 (73%)
	86..369	219/291 (74%)

Further analysis of the NOV38a protein yielded the following properties shown in
5 Table 38C.

Table 38C. Protein Sequence Properties NOV38a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome)
SignalP analysis:	Cleavage site between residues 70 and 71

A search of the NOV38a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 38D.

Table 38D. Geneseq Results for NOV38a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU10819	Human Type II GnRH-R splice variant 1 protein #2 - Homo sapiens, 377 aa. [WO200178796-A1, 25- OCT-2001]	40..323 26..322	278/297 (93%) 280/297 (93%)	e-156
AAU10816	Human Type II GnRH-R splice variant 3 protein - Homo sapiens, 366 aa. [WO200178796-A1, 25- OCT-2001]	40..323 15..311	278/297 (93%) 280/297 (93%)	e-156
AAU10815	Human Type II GnRH-R splice variant 2 protein - Homo sapiens, 376 aa. [WO200178796-A1, 25- OCT-2001]	40..323 25..321	278/297 (93%) 280/297 (93%)	e-156
AAU10814	Human Type II GnRH-R splice variant 1 protein #1 - Homo sapiens, 377 aa. [WO200178796-A1, 25- OCT-2001]	40..323 26..322	278/297 (93%) 280/297 (93%)	e-156
AAU10813	Human Type II GnRH-R - Homo sapiens, 379 aa. [WO200178796-A1, 25- OCT-2001]	40..323 28..324	278/297 (93%) 280/297 (93%)	e-156

In a BLAST search of public sequence databases, the NOV38a protein was found to have homology to the proteins shown in the BLASTP data in Table 38E.

Table 38E. Public BLASTP Results for NOV38a				
Protein Accession Number	Protein/Organism/Length	NOV38a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96P88	TYPE II GONADOTROPIN- RELEASING HORMONE RECEPTOR - Homo sapiens (Human), 379 aa (fragment).	40..323 28..324	281/297 (94%) 283/297 (94%)	e-158
CAD12280	SEQUENCE 9 FROM PATENT WO0178796 - Homo sapiens (Human), 366 aa.	40..323 15..311	278/297 (93%) 280/297 (93%)	e-156
CAD12279	SEQUENCE 7 FROM PATENT WO0178796 - Homo sapiens (Human), 376 aa.	40..323 25..321	278/297 (93%) 280/297 (93%)	e-156
CAD12278	SEQUENCE 5 FROM PATENT WO0178796 - Homo sapiens (Human), 377 aa.	40..323 26..322	278/297 (93%) 280/297 (93%)	e-156
CAD12277	SEQUENCE 3 FROM PATENT WO0178796 - Homo sapiens (Human), 379 aa.	40..323 28..324	278/297 (93%) 280/297 (93%)	e-156

PFam analysis predicts that the NOV38a protein contains the domains shown in the Table 38F.

Table 38F. Domain Analysis of NOV38a			
Pfam Domain	NOV38a Match Region	Identities/ Similarities for the Matched Region	Expect Value
7tm_1	66..319	60/287 (21%) 191/287 (67%)	9.7e-41

Example 39.

- 5 The NOV39 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 39A.

Table 39A. NOV39 Sequence Analysis			
	SEQ ID NO: 123	1047 bp	
NOV39a, CG96231-01 DNA Sequence	ATGTTTGGCCCCGCTAAAGGTCGCCATTTTGGAGTCCACCCGGCGCCTGGTTTCCCCG GCGGCGTCTCCCAACAGGCTGCCGGGACCAAGCTGGCCCCGCGGGTGCCTGGCCTGT GGGCAGCCGGACCGACACGATGTGGCGGCTCCGCTGCAAGGCCAAGGACGGCACCCTAT GTTTTGCAGGGGCTGTCCAGCCGGACCCGGGTGCGGGAATCCAGGGCCAAATTGCGG CCATCACCGGGATCGCCCCGGCGGTGAGCGAATCCTCGTCGGATACCCCTCCCGAGTG CCTGGATCTCAGCAATGGGGATACCATTTCTGGAAGACTTGCCCATCCAACTCTGGTGAC ATGCTGATCATTGAAGAAGACCAACCAGGCCCCAGAAGTTCACCTGCATTCTACTAAAC GTGGTGCTTCTAGTTACGTGAGGAACTTTGCCTGTGCTTACAGAACCGTGGTCCC AGCAGACAACCTTTGCCTCTTTACTAGTGTGTACTATGTGTCGAAGGAGGAGTCTTG AATCCAGCTTGTGCCCTGAGATGAGACGCTCATAGCACAAATTGTAGCAAGCGATC CAGACTTCTATAGTGAGGCAATACTGGGAAAAACAAATCAAGAGTACTGTGACTGGAT CAAAAGGGATGACACTTGGGGAGGAGCAATAGAGATATCGATTTTGTCCAAGTTTATC CAATGTGAAATATGTGTAGTGGATACACAGACAGTAAGAATTGATCGTTTGGGGGAG ATGCAGGATATACAAAAGGGTTCTGCTTATTTATGATGGCATCCACTATGATCCACT TCAGCGTAACTTCCCTGATCCAGATACACCTCCTCTGACCATTTTCTCCTCTAATGAT GATATTGTTCTTGTACAAGCACTGGAATTAGCAGATGAAGCTAGAAGAAGGAGACAGT TTACTGATGTCAACCGCTTCAACCTGAGATGCATGGTATGTGAGAAGGATTAACTGG ACAAGCAGAAGCAAGGGAACATGCCAAGGAGACAGGCCATACCAACTTTGGAGAAGTG TGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 1045	
	SEQ ID NO: 124	348 aa	MW at 38322.0kD
NOV39a, CG96231-01 Protein Sequence	MFGPAKGRHFGVHPAPGFPGGVSSQQAAGTKAGPAGAWPVGSRTDTMWRRLCKAKDGH VLQGLSSRTRVRELQGGIAAITGIAPGGQRIIVGYPPPECLDLSNGDTILEDLP IQSGD MLII EEDQTRPRSSPAFTKRGASSVYVRETLPLVLRTRTVVPADNSCLFTSVVYVVEGGVL NPACAPEMRRLIAQIVASDPDFYSEAILGKTNQBYCDWIKRDDTWGGAIEISILSKFY QCEICVVDQTQVRIDRFEGDAGYTKRVLLIYDGIHYDPLQRNFPDPPTPLTIFSSND DIVLVQALELADEARRRRQFTDVNRFTLRMCVCQKGLTGQAEAREHAKETGHTNFGVEV		
	SEQ ID NO: 125	1047 bp	
NOV39b, CG96231-02 DNA Sequence	ATGTTTGGCCCCGCTAAAGGTCGCCATTTTGGAGTCCACCCGGCGCCTGGTTTCCCCG GCGGCGTCTCCCAACAGGCTGCCGGGACCAAGCTGGCCCCGCGGTGCCTGGCCTGT GGGCAGCCGGACCGACACGATGTGGCGGCTCCGCTGCAAGGCCAAGGACGGCACCCTAT GTTTTGCAGGGGCTGTCCAGCCGGACCCGGGTGCGGGAATCCAGGGCCAAATTGCGG CCATCACCGGGATCGCCCCGGCGGTGAGCGAATCCTCGTCGGATACCCCTCCCGAGTG CCTGGATCTCAGCAATGGGGATACCATTTCTGGAAGACTTGCCCATCCAACTCTGGTGAC ATGCTGATCATTGAAGAAGACCAACCAGGCCCCAGAAGTTCACCTGCATTCTACTAAAC GTGGTGCTTCTAGTTACGTGAGGAACTTTGCCTGTGCTTACAGAACCGTGGTCCC AGCAGACAACCTTTGCCTCTTTACTAGTGTGTACTATGTGTCGAAGGAGGAGTCTTG AATCCAGCTTGTGCCCTGAGATGAGACGCTCATAGCACAAATTGTAGCAAGCGATC CAGACTTCTATAGTGAGGCAATACTGGGAAAAACAAATCAAGAGTACTGTGACTGGAT CAAAAGGGATGACACTTGGGGAGGAGCAATAGAGATATCGATTTTGTCCAAGTTTATC CAATGTGAAATATGTGTAGTGGATACACAGACAGTAAGAATTGATCGTTTGGGGGAG ATGCAGGATATACAAAAGGGTTCTGCTTATTTATGATGGCATCCACTATGATCCACT TCAGCGTAACTTCCCTGATCCAGATACACCTCCTCTGACCATTTTCTCCTCTAATGAT GATATTGTTCTTGTACAAGCACTGGAATTAGCAGATGAAGCTAGAAGAAGGAGACAGT TTACTGATGTCAACCGCTTCAACCTGAGATGCATGGTATGTGAGAAGGATTAACTGG ACAAGCAGAAGCAAGGGAACATGCCAAGGAGACAGGCCATACCAACTTTGGAGAAGTG TGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 1045	
	SEQ ID NO: 126	348 aa	MW at 38322.0kD
NOV39b, CG96231-02 Protein Sequence	MFGPAKGRHFGVHPAPGFPGGVSSQQAAGTKAGPAGAWPVGSRTDTMWRRLCKAKDGH VLQGLSSRTRVRELQGGIAAITGIAPGGQRIIVGYPPPECLDLSNGDTILEDLP IQSGD MLII EEDQTRPRSSPAFTKRGASSVYVRETLPLVLRTRTVVPADNSCLFTSVVYVVEGGVL NPACAPEMRRLIAQIVASDPDFYSEAILGKTNQBYCDWIKRDDTWGGAIEISILSKFY QCEICVVDQTQVRIDRFEGDAGYTKRVLLIYDGIHYDPLQRNFPDPPTPLTIFSSND DIVLVQALELADEARRRRQFTDVNRFTLRMCVCQKGLTGQAEAREHAKETGHTNFGVEV		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 39B.

Table 39B. Comparison of NOV39a against NOV39b.		
Protein Sequence	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV39b	1..348 1..348	344/348 (98%) 344/348 (98%)

Further analysis of the NOV39a protein yielded the following properties shown in Table 39C.

Table 39C. Protein Sequence Properties NOV39a	
PSort analysis:	0.3000 probability located in nucleus; 0.2271 probability located in lysosome (lumen); 0.1109 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV39a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 39D.

Table 39D. Geneseq Results for NOV39a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB60807	Drosophila melanogaster polypeptide SEQ ID NO 9213 - Drosophila melanogaster, 347 aa. [WO200171042-A2, 27-SEP-2001]	49..348 7..347 .	144/343 (41%) 196/343 (56%)	2e-64
AAG53575	Arabidopsis thaliana protein fragment SEQ ID NO: 68223 - Arabidopsis thaliana, 146 aa. [EP1033405-A2, 06-SEP- 2000]	148..271 4..124	63/124 (50%) 87/124 (69%)	8e-31
AAG20460	Arabidopsis thaliana protein fragment SEQ ID NO: 22659 - Arabidopsis thaliana, 146 aa. [EP1033405-A2, 06-SEP- 2000]	148..271 4..124	63/124 (50%) 87/124 (69%)	8e-31
AAG53576	Arabidopsis thaliana protein fragment SEQ ID NO: 68224 - Arabidopsis thaliana, 122 aa. [EP1033405-A2, 06-SEP- 2000]	179..271 8..100	52/93 (55%) 67/93 (71%)	1e-24
AAG20461	Arabidopsis thaliana protein fragment SEQ ID NO: 22660 - Arabidopsis thaliana, 122 aa. [EP1033405-A2, 06-SEP- 2000]	179..271 8..100	52/93 (55%) 67/93 (71%)	1e-24

In a BLAST search of public sequence databases, the NOV39a protein was found to have homology to the proteins shown in the BLASTP data in Table 39E.

Table 39E. Public BLASTP Results for NOV39a				
Protein Accession Number	Protein/Organism/Length	NOV39a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
AAL39440	GM14814P - Drosophila melanogaster (Fruit fly), 347 aa.	49..348 7..347	144/343 (41%) 196/343 (56%)	5e-64
Q9VRJ9	CG4603 PROTEIN - Drosophila melanogaster (Fruit fly), 347 aa.	49..348 7..347	144/343 (41%) 196/343 (56%)	5e-64
Q9LPT6	F11F12.1 PROTEIN (HYPOTHETICAL 23.4 KDA PROTEIN) - Arabidopsis thaliana (Mouse-ear cress), 208 aa.	148..347 4..206	96/206 (46%) 131/206 (62%)	1e-44
O13974	HYPOTHETICAL 35.7 KDA PROTEIN C24C9.14 IN CHROMOSOME 1 - Schizosaccharomyces pombe (Fission yeast), 329 aa.	48..347 5..328	101/333 (30%) 161/333 (48%)	4e-36
P43558	Hypothetical 33.5 kDa protein in SEC53-FET5 intergenic region - Saccharomyces cerevisiae (Baker's yeast), 301 aa.	92..347 39..299	90/278 (32%) 144/278 (51%)	7e-33

PFam analysis predicts that the NOV39a protein contains the domains shown in the Table 39F.

Table 39F. Domain Analysis of NOV39a			
Pfam Domain	NOV39a Match Region	Identities/ Similarities for the Matched Region	Expect Value
OTU	154..269	40/135 (30%) 87/135 (64%)	1.1e-18

Example 40.

- 5 The NOV40 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 40A.

Table 40A. NOV40 Sequence Analysis			
	SEQ ID NO: 127	591 bp	
NOV40a, CG96260-01 DNA Sequence	AACAGAGGTTTCAGCTGGGAGACCCCTCACACACAGGTGAAGGCGTGGCTGTAGAACCT CAATGGACCCCTGGTCTCCTCAGGAATGAAGGTCATTGCCATCCTCACCTCCTCCT CTTCTGCTCCCCACCCACAGCAGTTTCTGGCAGTTTCAGAGGAGGGTCAAACACATC ACGGGGCGAAGTGCCTTCTTCTCATATTACGGATATGGCTGCTACTGTGGGCTTGGGG ATAAAGGGATCCCCGTGGATGACACTGACAGGCACAGCCCTCATCTCCCTCTCCCTA CGAGAAGCTGAAGGAGTTTCAGCTGCCAGCCTGTGTTGAACAGCTACCAGTTCACATC GTCAATGGCGCAGTGGTAGGTGGATGCACCCTTGGTCCTGGTGCCAGCTGCCACTGCA GGCTGAAGGCCTGTGAGTGTGACAAGCAATCCGTGCACTGCTTCAAAGAGAGCCTGCC CACCTATGAGAAAACTTCAAGCAGTTCTCCAGCCAGCCAGGTGTGGCAGACATAAG CCCTGGTGCTAGGGACACCACAGGGTCCCTCTCATCATCCAGCATCCGCTCTAGTGT GCTCTTCCAGG		
	ORF Start: ATG at 61	ORF Stop: TAG at 532	
	SEQ ID NO: 128	157 aa	MW at 17585.0kD
NOV40a, CG96260-01 Protein Sequence	MDPLVSSGMKVIAILTLFLFCSPHSSFWQFORRVKHITGRSAFFSYGYGICYCGLGD KGIPVDDTDRHSPSSPSPYEKLKEFSCQPVLSYQFHI VNGAVVGGCTLGPGASCHCR LKACECDKQSVHCFKESLPTYEKNFKQFSSQPRCGRHKPWC		

Further analysis of the NOV40a protein yielded the following properties shown in Table 40B.

Table 40B. Protein Sequence Properties NOV40a	
PSort analysis:	0.8200 probability located in endoplasmic reticulum (membrane); 0.2222 probability located in microbody (peroxisome); 0.1900 probability located in plasma membrane; 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 27 and 28

A search of the NOV40a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 40C.

Table 40C. Geneseq Results for NOV40a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV40a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG63223	Amino acid sequence of a human lipid metabolism enzyme - Homo sapiens, 152 aa. [WO200153468-A2, 26-JUL-2001]	11..157 2..152	139/151 (92%) 140/151 (92%)	1e-82
AAR63044	RPLA2-8 - Rattus sp, 158 aa. [WO9502328-A, 26-JAN-1995]	1..157 1..158	106/160 (66%) 121/160 (75%)	1e-55
AAB12536	Mouse secretory phospholipase A2 protein sequence SEQ ID NO:14 - Mus musculus, 144 aa. [WO200034486-A1, 15-JUN-2000]	17..157 10..144	48/141 (34%) 70/141 (49%)	1e-17
AAB11994	Mouse secreted phospholipase A2 - Mus musculus, 144 aa. [JP2000166544-A, 20-JUN-2000]	17..157 10..144	48/141 (34%) 70/141 (49%)	1e-17
AAB56432	Human prostate cancer antigen protein sequence SEQ ID NO:1010 - Homo sapiens, 164 aa. [WO200055174-A1, 21-SEP-2000]	9..150 21..164	48/150 (32%) 72/150 (48%)	2e-16

In a BLAST search of public sequence databases, the NOV40a protein was found to have homology to the proteins shown in the BLASTP data in Table 40D.

Table 40D. Public BLASTP Results for NOV40a				
Protein Accession Number	Protein/Organism/Length	NOV40a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
B54762	phospholipase A2 (EC 3.1.1.4) RPLA2-8 precursor - rat, 158 aa.	1..157 1..158	106/160 (66%) 121/160 (75%)	2e-55
P39878	Group IIC secretory phospholipase A2 precursor (EC 3.1.1.4) (Phosphatidylcholine 2- acylhydrolase GIIC) (GIIC sPLA2) (PLA2-8) (14 kDa phospholipase A2) - Rattus norvegicus (Rat), 150 aa.	9..157 1..150	99/152 (65%) 114/152 (74%)	8e-52
P48076	Group IIC secretory phospholipase A2 precursor (EC 3.1.1.4) (Phosphatidylcholine 2- acylhydrolase GIIC) (GIIC sPLA2) (PLA2-8) (14 kDa phospholipase A2) - Mus musculus (Mouse), 150 aa.	9..157 1..150	98/152 (64%) 114/152 (74%)	9e-51
A54762	phospholipase A2 (EC 3.1.1.4) MPL2-8 - mouse, 130 aa (fragment).	27..157 1..130	90/132 (68%) 103/132 (77%)	5e-48
A60718	phospholipase A2 homolog, non- pancreatic - human, 50 aa (fragment).	22..68 4..50	46/47 (97%) 47/47 (99%)	6e-23

PFam analysis predicts that the NOV40a protein contains the domains shown in the Table 40E.

Table 40E. Domain Analysis of NOV40a			
Pfam Domain	NOV40a Match Region	Identities/ Similarities for the Matched Region	Expect Value
phoslip	27..68	19/43 (44%) 31/43 (72%)	1.7e-13
phoslip	77..150	20/82 (24%) 52/82 (63%)	4.2e-06

Example 41.

- 5 The NOV41 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 41A.

Table 41A. NOV41 Sequence Analysis			
	SEQ ID NO: 129	952 bp	
NOV41a, CG96364-01 DNA Sequence	CCTGCCGCTGCCTCCTCTTTCTCCTTTCAACATGACAGATGCCGCTGTGTCTTCCACCA AGGACTTCCTGGCAGGTGGAGTGGCAGCAGCCATCTCCAAGATGGGGGTAGCGCCCA CGAGAGGGTCAAGCTGCTGCTGGAGGTACAGCATGCCAGCAAGCAGATCACCGCAGAT ATGCAATACAAGGGCATTATAGACTGCGTGGTCCATATTCTCAAGGAGCAGGGAGTCC TGTCTTCTGGCACCGTAACCTGGTCAATGTCTCTCAGATACTTCCCCACCCAGGCTCT CAACTTGGCCTTCAAAGATAAACACAAGCAGATCTTCTCTGGGTGGTGTGGACAAGAG ACCCAGTTTGGGAGCTACTTTGCAGGGAATCCGGCATCAGGTGGTGCCGCTGGGGCCA CATCCTTGTGTTTTTGTGTACCCTCTTGATTTTGCCCATACCTGTCTAGCAGCTGACG GGGTAAAGCTGGAGCTGAAAGGGAATTCCGAGGCCTCGGTAACCTGCTGGTTAAGATC TACGAACTGATGGGATTAAGGGCCTGTACCAAGGCTTTAACGTGTCTGTGCAGGGT TTATCATCTACCGAGCTGCCTACTTTGGTATCTATGACACTGCAAGGGAATGCTTCC GGATCCCAGGAACACTCACATCATCATCAGCTGGATGATCACTCAGACTGTCTACTGCC ATTGCTGGGTGACTTCCCATTCCATTTGACACCGTTTCGATGCCGATGATGATGCAGT CAGGGCGCAAAGGAATTGACGTCATGTACACAGGCACGCTTGACTGCTGGAGGAAGAT TGTGGTGTGAAGGAGGCAAAGATTTTTCAAGGGTGCATGGTCCAGTGTCTCAGAG GGCATGGATGGTGTCTTTGTGCTTGTCTTGTATGGTGAAATCAAGAAGTACACATAAG TTATTTCTCTGGGATTTCCCCCACT		
	ORF Start: ATG at 31	ORF Stop: TAA at 925	
	SEQ ID NO: 130	298 aa	MW at 32668.6kD
NOV41a, CG96364-01 Protein Sequence	MTDAAVSFTKDFLAGGVAAAI SKMGVAPIERVKLLLEVQHASKQITADMQYKGI IDC VHILKEQGVLSFWHRNLNVNLRYPFTQALNLAFAKDKHKQIFLGGVDKRTQFGSYFAG PASGGAAGATSLCFVYPLDFAHTCLAADVKGAGAEREFRGLGNCLVKIYETDGIKGL QGFNVSVQGI I IYRAAYFGIYDTAKGMLPDRNTHI I ISWITQVTAIAGLTSHPFD TVRCRMMMQSGRKGIDVMTYGTLDLCWRKIVGDEGGKDFFKGAWSSVLRGMDGAFVLV YGEIKKYT		
	SEQ ID NO: 131	745 bp	
NOV41b, CG96364-03 DNA Sequence	CCGCAGCGCCGTAGTCAAACCGAACCCGCGCCAGTCCCGTCTGCAGCAGTCTGCCTC CTTCTTTTCAACATGACAGATGCCGCTATTGCTTTCGCAAGGACTTCTCTGGCAGGTG AGTGGCCGCGAGCCATCTCCAAGACGGCGGTAGCGCCCATCGAGCGGGTCAAGCTGCT CTGCAGGTGCAGCATGCCAGCAAGCAGATCACTGCAGATAAGCAATACAAGGCATT TAGACTGCGTGGTCCGTATTCCCAAGGAGCAGGAAGTTCTGTCTTCTTGGCGCGGTA CCTGGCCAATGTCTCAGATACTTCCCCACCCAGGCTCTTAACCTTCGCTTCAAGAAT AAATACAAGCAGATCTTCTGGGTGGTGTGGACAAGAGAACCAGTTTGGCGCTACT TTGCAGGGAATCTGGCATCGGGTGGTGGCGCAGGGGCCACATCCCTGTGTTTTGTGT CCCTCTTGATTTTGCCCGTACCCGTCTAGCAGCTGATGTGGGTAAAGCTGGAGCTGA AGGGAATTCGAGGCCTCGGTGACTGCTGTTAAGATCTACAATCTGATGGGATT AGGGCCTGTACCAAGGCCTTGACTGCTGGCGGAAGATTGCTCGTGATGAAGGAGGCA AGCTTTTTTCAAGGGTGCATGGTCCAATGTTCTCAGAGGCATGGGTGGTGTCTTTGT CTTGTCTTGTATGATGAAATCAAGAAGTACACATAAGTTATTTCCTAGG		
	ORF Start: ATG at 70	ORF Stop: TAA at 730	
	SEQ ID NO: 132	220 aa	MW at 24272.9kD
NOV41b, CG96364-03 Protein Sequence	MTDAALSFAKDFLAGGVAAAI SKTAVAPIERVKLLQVQHASKQITADKQYKGI IDC VRIPKEQEVLSFWRGNLANVIRYPFTQALNLAFAKDKYKQIFLGGVDKRTQFWRYFAG LASGGAAGATSLCFVYPLDFARTRLAADVKGAGAEREFRGLGDCLVKIYKSDGIKGL QGLDCWRKIARDEGGKAFKGAWSNVLRGMGGAFVLVLYDEIKKYT		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 41B.

Table 41B. Comparison of NOV41a against NOV41b.		
Protein Sequence	NOV41a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV41b	1..178	142/178 (79%)
	1..178	150/178 (83%)

Further analysis of the NOV41a protein yielded the following properties shown in Table 41C.

Table 41C. Protein Sequence Properties NOV41a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.3600 probability located in mitochondrial matrix space; 0.3000 probability located in mitochondrial intermembrane space; 0.2883 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV41a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 41D.

Table 41D. Geneseq Results for NOV41a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV41a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU10379	Human adenine nucleotide translocator 2 (ANT2) - Homo sapiens, 298 aa. [WO200185944-A2, 15-NOV-2001]	1..298 1..298	261/298 (87%) 274/298 (91%)	e-150
AAU01199	Human adenine nucleotide translocator-2 (ANT-2) protein - Homo sapiens, 298 aa. [WO200132876-A2, 10-MAY-2001]	1..298 1..298	261/298 (87%) 274/298 (91%)	e-150
AAY71032	Human adenine nucleotide translocator ANT2 - Homo sapiens, 298 aa. [WO200026370-A2, 11-MAY-2000]	1..298 1..298	261/298 (87%) 274/298 (91%)	e-150
AAU10380	Human adenine nucleotide translocator 3 (ANT3) - Homo sapiens, 298 aa. [WO200185944-A2, 15-NOV-2001]	1..296 1..296	242/296 (81%) 267/296 (89%)	e-143
AAU01200	Human adenine nucleotide translocator-3 (ANT-3) protein - Homo sapiens, 298 aa. [WO200132876-A2, 10-MAY-2001]	1..296 1..296	242/296 (81%) 267/296 (89%)	e-143

In a BLAST search of public sequence databases, the NOV41a protein was found to have homology to the proteins shown in the BLASTP data in Table 41E.

Table 41E. Public BLASTP Results for NOV41a				
Protein Accession Number	Protein/Organism/Length	NOV41a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P05141	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Homo sapiens (Human), 298 aa.	1..298 1..298	263/298 (88%) 275/298 (92%)	e-151
Q09073	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Rattus norvegicus (Rat), 298 aa.	1..298 1..298	261/298 (87%) 274/298 (91%)	e-150
A29132	ADP,ATP carrier protein T2 - human, 298 aa.	1..298 1..298	261/298 (87%) 274/298 (91%)	e-150
P51881	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Mus musculus (Mouse), 298 aa.	1..298 1..298	260/298 (87%) 273/298 (91%)	e-149
BAB84673	ADENINE NUCLEOTIDE TRANSLOCATOR 2 - Bos taurus (Bovine), 298 aa.	1..298 1..298	259/298 (86%) 273/298 (90%)	e-149

PFam analysis predicts that the NOV41a protein contains the domains shown in the Table 41F.

Table 41F. Domain Analysis of NOV41a			
Pfam Domain	NOV41a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	7..105	34/125 (27%) 85/125 (68%)	4.6e-22
mito_carr	112..208	35/125 (28%) 82/125 (66%)	7.4e-22
mito_carr	209..298	27/125 (22%) 66/125 (53%)	3.2e-09

Example 42.

The NOV42 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 42A.

Table 42A. NOV42 Sequence Analysis			
	SEQ ID NO: 133	936 bp	
NOV42a, CG96422-01 DNA Sequence	GCCACCATGACGGAACAAGCCATCTCCTTCGCCAAGGACTTCCTAGCTGGAGGCATCA CCGCCGCTATCTCCAAGACAGCCGTGGCATCGATCAAGCGGGTCCAGCTGCTGCTGCA GATGCAGCACGCCAGCATGCCGATGGCCGCCGCCAAGCAGTGC AAGGGGCATCGTGGAG TGCACTGTCGGCATCCCCAAGATCAGGGCGTGCTGTCTTCTGGAGGGGCAACCTGG CCAATGTCATCCGCTACTCCCCACGCAAGCCCTCAATTTGCGCTTCAAGGATAAGTA CAAGCAGATCTTCTTGGCGGGCGTGGACAAGCACACGCAGTTCTGCAGGTACTTTGCG GGCAACCTGGCCCTCTGGCGGCACGGCCGTGCTGTATCCCTGGATTTCACCAGAACCC GCCTGGCAGCCGACGTGGGAAAGTCAGGCACGGAGCGGAGTTCGAGGCCTGGGAGA CTGCCTAGTGAAGATCAGCAAGTCCGACGGCATCCGGGGCCTTTACCAGGGCTTCAG GTCTCCGTGCAGGCCATCATCATCTACCAGGCAGCCTACTTCAGGGTGTACGATACGG CCAATGGCATGTTCCCGACCCCAAGAACACACATCTCTGGTGAGCTGGATGACCGC GCAGACCGTGACGGCCGTGGCTGGCGTGCTCTCCAACCCCTTCGACACAGTGC GCGCG CGGACGATGATGCAGTCCCGGCGCAAAGGAGCTGACATCATGTACACGGGCACCGTCC ACTGTTGGAGGAAGATCTTCAGAGATGAGCGGGGCGAGGCCTTTTCAAGGGTGTGTG GTCCAAGCCCTCAAAGGCATGGGGGTGGGGGCGGGCTTCGTGCTGGTCTGTACGAC GAGCTGAAGGTGAGCTAAGGGCCTTGGCATCTCCACACACACCAAGGAACCAAG AGAACCAC		
	ORF Start: ATG at 7	ORF Stop: TAA at 886	
	SEQ ID NO: 134	293 aa	MW at 32354.3kD
NOV42a, CG96422-01 Protein Sequence	MTEQAISFAKDFLAGGITAAISKTAVASIKRVQLLLQMQHASMPMAAAKQCKGIVDCI VRIPKDDQVLSFWRGNLANVIRYSPTQALNFAFKDKYKQIFLAGVDKHTQFCRYFAGN LASGGTAVVYPLDFTRTLAADVGKSGTEREFRGLGDCLVKISKSDGIRGLYQGFVS VQAII IYQAAYFRVYDTANGMFPDPKNTHILVSWMTAQTVTAVAGVLSKPFDTVRRRT MMQSRRKADIMYTGTVDCWRKIFRDERGEAFFKGVWSNALKMGVAGFVLVLYDEI KVS		

Further analysis of the NOV42a protein yielded the following properties shown in
5 Table 42B.

Table 42B. Protein Sequence Properties NOV42a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.2809 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV42a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 42C.

Table 42C. Geneseq Results for NOV42a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV42a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU10380	Human adenine nucleotide translocator 3 (ANT3) - Homo sapiens, 298 aa. [WO200185944-A2, 15-NOV-2001]	1..291 1..295	253/297 (85%) 264/297 (88%)	e-140
AAU01200	Human adenine nucleotide translocator-3 (ANT-3) protein - Homo sapiens, 298 aa. [WO200132876-A2, 10-MAY-2001]	1..291 1..295	253/297 (85%) 264/297 (88%)	e-140
AAM41427	Human polypeptide SEQ ID NO 6358 - Homo sapiens, 323 aa. [WO200153312-A1, 26-JUL-2001]	1..291 26..320	253/297 (85%) 264/297 (88%)	e-140
AAM39641	Human polypeptide SEQ ID NO 2786 - Homo sapiens, 298 aa. [WO200153312-A1, 26-JUL-2001]	1..291 1..295	253/297 (85%) 264/297 (88%)	e-140
AAY71033	Human adenine nucleotide translocator ANT3 - Homo sapiens, 298 aa. [WO200026370-A2, 11-MAY-2000]	1..291 1..295	253/297 (85%) 264/297 (88%)	e-140

In a BLAST search of public sequence databases, the NOV42a protein was found to have homology to the proteins shown in the BLASTP data in Table 42D.

Table 42D. Public BLASTP Results for NOV42a

Protein Accession Number	Protein/Organism/Length	NOV42a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P12236	ADP,ATP carrier protein, liver isoform T2 (ADP/ATP translocase 3) (Adenine nucleotide translocator 3) (ANT 3) - Homo sapiens (Human), 298 aa.	1..291 1..295	253/297 (85%) 264/297 (88%)	e-140
P32007	ADP,ATP carrier protein, isoform T2 (ADP/ATP translocase 3) (Adenine nucleotide translocator 3) (ANT 3) - Bos taurus (Bovine), 298 aa.	1..291 1..295	246/297 (82%) 261/297 (87%)	e-136
P05141	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Homo sapiens (Human), 298 aa.	1..291 1..295	234/297 (78%) 256/297 (85%)	e-132
BAB84673	ADENINE NUCLEOTIDE TRANSLOCATOR 2 - Bos taurus (Bovine), 298 aa.	1..291 1..295	233/297 (78%) 254/297 (85%)	e-131
Q09073	ADP,ATP carrier protein, fibroblast isoform (ADP/ATP translocase 2) (Adenine nucleotide translocator 2) (ANT 2) - Rattus norvegicus (Rat), 298 aa.	1..291 1..295	232/297 (78%) 255/297 (85%)	e-131

PFam analysis predicts that the NOV42a protein contains the domains shown in the Table 42E.

Table 42E. Domain Analysis of NOV42a			
Pfam Domain	NOV42a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	7..105	33/125 (26%) 83/125 (66%)	9.2e-23
mito_carr	112..202	38/125 (30%) 78/125 (62%)	1.9e-18
mito_carr	203..293	25/125 (20%) 69/125 (55%)	2.9e-05

Example 43.

The NOV43 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 43A.

Table 43A. NOV43 Sequence Analysis			
	SEQ ID NO: 135	858 bp	
NOV43a, CG96442-01 DNA Sequence	ATGTTATGGGTATCTTTGATCTCTGAATTTGAAGGAGGGCTTACTGAGTGTGTCTGGC TCATTGGCACTCCTGAGGCAGAACACTTTTCAGGCAAGGGAGTACTACCTGTGCTGTT CAAATTACGTCCTGTGGGTGGCCAACTACCTCATCTGAATGCAGGGCTTGTGGTGGTT TGGTATCGGAACATCATCATCCTTTCCAACCTCCATAGGGCACCTTCCAGAACAC AGAAACTGGAAGAAAAATGCTTGCTCTAGGGCTCCTTGGGCAGTGGCTCAGTGGCAT CTGCCTCAGTGGCTCCCCTTTTCCATACTCCTGGGAAAAGTTTGTCTTGGAAATGAT GGGTATTTAGAAGTGGCCTACTTTTCATGTCAAGGAAGAAAGTAAGAGAGAGAGAGAG GAGAGGAGAACGGAGAACATTCTCCCATCAAAATGTCCCTTCCCAAATTCATCAGT GCCTCTTCTAGGCTTTGAGAGCATGTACTGTGTACACGCTGCTATCACTGCATTTTCAT CCCCTATGGCTGTCCCTTCACTGCACGGGGCAGCCCTGTGCTTCCGTTTTTCAAGA TCGAATGTAGCAAACAAGTTATCTATGCAGTTAATGATGATTTATTGAGGGCCTACAA AACGGGCACAGTGGCAACGGCTTCGGGGAAGAAAAGAGGGCAAGGTGCACGGGAAGAA ATGAGCTCAAAGGAGATTGTGCCAATTACTGCACCTGTGAGTGTCAACAACAGATAA AATTAATTTCGGAACCCAAACAGTCTGCAGAAAACCGAACACAATTATCTAAGAAATT AAGTTATACCGTTGCCATGACCTTTAATCAATCATCCAACCGTTAA		
	ORF Start: ATG at 1	ORF Stop: TAA at 856	
	SEQ ID NO: 136	285 aa	MW at 31776.3kD
NOV43a, CG96442-01 Protein Sequence	MLVWSLISEFEGGLTECVWLIGTPEAEHFSGKGVLVPLFKLRVLVGGQLPHLNAGLVV WYRNSSSFPTSIGHLPEQQKLERKMLALGLLGQNLSGICLSGSPFPYSWEKFLVGNL GYLEVAYFHVKEESKREREREENGEHSPHQNVPSQILSVPLLGFESEMYCVHAAITAFH PLWLSPSLHGAALCFPPFKIECSKQVIYAVNDLLRAYKTGTATASGKKRGQGAAREE MSSKEIVPITAPVSVQQIKLRKPKQSAENRTQLSKKLSYTVAMTFNQSSNR		

Further analysis of the NOV43a protein yielded the following properties shown in Table 43B.

Table 43B. Protein Sequence Properties NOV43a	
PSort analysis:	0.4010 probability located in microbody (peroxisome); 0.3000 probability located in nucleus; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	Cleavage site between residues 16 and 17

A search of the NOV43a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 43C.

Table 43C. Geneseq Results for NOV43a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV43a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG01918	Novel human diagnostic protein #1909 - Homo sapiens, 434 aa. [WO200175067-A2, 11-OCT-2001]	59..140 219..314	30/100 (30%) 41/100 (41%)	5.8
ABG01918	Novel human diagnostic protein #1909 - Homo sapiens, 434 aa. [WO200175067-A2, 11-OCT-2001]	59..140 219..314	30/100 (30%) 41/100 (41%)	5.8
ABG26879	Novel human diagnostic protein #26870 - Homo sapiens, 800 aa. [WO200175067-A2, 11-OCT-2001]	91..156 328..393	20/68 (29%) 31/68 (45%)	7.6
ABG26879	Novel human diagnostic protein #26870 - Homo sapiens, 800 aa. [WO200175067-A2, 11-OCT-2001]	91..156 328..393	20/68 (29%) 31/68 (45%)	7.6

AAY08899	S. peregrina cysteine protease-like protein 1 - Sarcophaga peregrina, 550 aa. [JP11146789-A, 02-JUN-1999]	66..111 286..327	17/46 (36%) 25/46 (53%)	9.9
----------	---	---------------------	----------------------------	-----

In a BLAST search of public sequence databases, the NOV43a protein was found to have homology to the proteins shown in the BLASTP data in Table 43D.

Table 43D. Public BLASTP Results for NOV43a				
Protein Accession Number	Protein/Organism/Length	NOV43a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9X935	CONSERVED HYPOTHETICAL PROTEIN - Streptomyces coelicolor, 309 aa.	131..192 192..248	19/62 (30%) 27/62 (42%)	8.4
Q9RNK7	HYPOTHETICAL PROTEIN - Zymomonas mobilis, 415 aa.	141..178 221..258	12/38 (31%) 24/38 (62%)	8.4

PFam analysis predicts that the NOV43a protein contains the domains shown in the Table 43E.

Table 43E. Domain Analysis of NOV43a			
Pfam Domain	NOV43a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

5 Example 44.

The NOV44 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 44A.

Table 44A. NOV44 Sequence Analysis		
	SEQ ID NO: 137	999 bp

NOV44a, CG96501-01 DNA Sequence	GGTTTAAATGCTTCCGGGTTGCCACTGCAGTGGCATTCTGACTGTGGGAGCCTCAGT TTCCAGTCATTGATGAGCCTGAGCAGCAGCAAAGTACCACTGTAAGCCATGAGATGT CTCGTCTGAAGTGGAAACCCCTTTGTATATGACGCCCTTGCCCTCTACTGCTGAGTT TGGAACCTTCCCATGGACCTTGCCAAAACACGACTTCAGGTACAAGGCCAAAGCATT GATGTCCGTTTCAAAGAAACAAAATATAGACGGATGTTTCATGCTTTGTTTGGATCT ATAAAGCGGAGGGGGGATTGGCTCTGTATTAGGAATTGCTCCTGTTTGCAAAGACA AGCATCATATGGCACCATTAAAATTGGGATTACCAAAGCTTGAAGCAATTATCTGTA GAACGTTTAGAAGATGAACTCTTTAATCAACATGATCTGTGGGTTAGTGTGAGGGG TGATATTTTCCACTATAGCCAATCCACCGATGTTCTAAAGATTGGAATGCAGGCTCA AGGAAGTTTGTTCAGGGGAGCATGATTGGCAGCTTCATCGATATATACCAACAGGAA GGCACCAGGGGTCTGTGGAGGAGTGTGGTTCCAAGTCTCAGCATGCTGCCATCGTTG TGGGAGTAGAGCTACCACTCTATGATTTTACTAAGAAGCACTTAATATTGTGAGGAAT GATGGAAGACACAACCTTAACTCACTTTGTTCCAGCTTTACATATGTTTGGCTGGG GCTCTTGCTCTAATCCAGGTGATGTGGCAGGCACTCACGTGATGAACAGAGGGCAA TCGTGGGACATGTGATCTCTATAAGGGCACTTTGGATGGTATTTTAAAAATGTGGAA ACATGAGGGCTTTTTTTGTATTCTAAAGATTTTGGCCAACTGGCTTTGGCGTGGAA CCCTGGAACATCATCTTAAATTACATATGAGAAGCTCAAGAGGCTTAAATCTAAG GACTGAATTATAT
------------------------------------	--

	ORF Start: ATG at 8	ORF Stop: TAA at 977	
	SEQ ID NO: 138	323 aa	MW at 36200.6kD
NOV44a, CG96501-01 Protein Sequence	MLPGCHCSGISDCGSLSFVIDEPEQQQSTTVSHEMSRLNWKPFVYDALASITAETFGT FPMDLAKTRLQVQGQSIDVRFKETKYRRMPHALFWIYKAEGGLALYSIGIAPVLQRQAS YGTIKIGIYQSLKQLSVLEDETLLINMICGVVSGVIFSTIANPTDVLKIRMQAQGS LFQGS MIGSFIDYQQEGRGLWRSVVPTAQHAAIVVGVELPVYDFTKKHLILSGMME DTTLTHFVSSFTYGLAGALASNP GDVAGTHVMNQRAIVGHVDLYKGTLDGILMKWKHE GFFLYSKGFWPWNLWRGPWNIILKITYEKLRKRL		

Further analysis of the NOV44a protein yielded the following properties shown in Table 44B.

Table 44B. Protein Sequence Properties NOV44a	
PSort analysis:	0.7480 probability located in microbody (peroxisome); 0.7000 probability located in plasma membrane; 0.2000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV44a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 44C.

Table 44C. Geneseq Results for NOV44a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV44a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY94669	Murine uncoupling protein isoform mUCP5S amino acid sequence - Mus musculus, 325 aa. [WO200032624-A2, 08- JUN-2000]	27..323 27..323	257/297 (86%) 266/297 (89%)	e-144
AAY94668	Murine uncoupling protein isoform mUCP5L amino acid sequence - Mus musculus, 322 aa. [WO200032624-A2, 08- JUN-2000]	27..323 24..320	257/297 (86%) 266/297 (89%)	e-144
AAE06056	Human gene 16 encoded secreted protein HMIAP86, SEQ ID NO:118 - Homo sapiens, 334 aa. [WO200151504-A1, 19- JUL-2001]	27..323 36..332	258/297 (86%) 264/297 (88%)	e-144
AAY87079	Human secreted protein sequence SEQ ID NO:118 - Homo sapiens, 335 aa. [WO200004140-A1, 27- JAN-2000]	27..323 36..332	258/297 (86%) 264/297 (88%)	e-144
AAY94666	Human uncoupling protein isoform hUCP5S amino acid sequence - Homo sapiens, 322 aa. [WO200032624-A2, 08- JUN-2000]	27..323 24..320	258/297 (86%) 264/297 (88%)	e-144

In a BLAST search of public sequence databases, the NOV44a protein was found to have homology to the proteins shown in the BLASTP data in Table 44D.

Table 44D. Public BLASTP Results for NOV44a

Protein Accession Number	Protein/Organism/Length	NOV44a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9Z2B2	Brain mitochondrial carrier protein-1 (BMCP-1) (Mitochondrial uncoupling protein 5) (UCP 5) (Solute carrier family 25, member 14) - Mus musculus (Mouse), 325 aa.	27..323 27..323	257/297 (86%) 266/297 (89%)	e-144
O95258	Brain mitochondrial carrier protein-1 (BMCP-1) (Mitochondrial uncoupling protein 5) (UCP 5) (Solute carrier family 25, member 14) - Homo sapiens (Human), 325 aa.	27..323 27..323	258/297 (86%) 264/297 (88%)	e-144
Q9JMH0	BRAIN MITOCHONDRIAL CARRIER PROTEIN-1 - Rattus norvegicus (Rat), 322 aa.	27..323 24..320	255/297 (85%) 265/297 (88%)	e-143
Q9EP88	BRAIN MITOCHONDRIAL CARRIER PROTEIN BMCP1 (BRAIN MITOCHONDRIAL CARRIER PROTEIN-1) - Rattus norvegicus (Rat), 325 aa.	27..323 27..323	255/297 (85%) 265/297 (88%)	e-143
Q9CR58	4933433D23RIK PROTEIN - Mus musculus (Mouse), 291 aa.	36..323 1..289	205/289 (70%) 231/289 (78%)	e-113

PFam analysis predicts that the NOV44a protein contains the domains shown in the Table 44E.

Table 44E. Domain Analysis of NOV44a			
Pfam Domain	NOV44a Match Region	Identities/ Similarities for the Matched Region	Expect Value
mito_carr	39..138	36/125 (29%) 75/125 (60%)	1.5e-17
mito_carr	140..231	27/125 (22%) 74/125 (59%)	2.8e-21
mito_carr	234..323	29/125 (23%) 70/125 (56%)	0.00021

Example 45.

The NOV45 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 45A.

Table 45A. NOV45 Sequence Analysis			
	SEQ ID NO: 139	1065 bp	
NOV45a, CG96557-01 DNA Sequence	ATGCTCCGCAGACTGGCCCCGAGCGGCGGCCGTGGTGGGCCTGCGCAGCGACAAGGTG CCAGGCCCGGGCCGCGCCATGAGGTGCGTCATGGCCGAGGCGGTCTGTGGGGACGACGA CTACGGCGAAGACCGGGTCCGTGATGAGCTGCAAGAAATGGCTGTGGAGCTGCTAGGG CTGGGGCGCACCTCTTTGTGCCCTCCAACACTATGGCCAACTCATCTCTGGGTCTG ACGACTGCAGGTGCCCTCGGCTGCGTGTCCACCAAGAAGCAGCATTGCAGCAGCTTCTC CAGCAGGCGATGCTCTGTGGAGGGGCTCCAGGCGCAGAGGGCTGGAGTGCACCTCC CTCCCTCCCAGAGCTGCCCCACTGCACCTCTGGACTGGCTGAGCTGGAGAGGATGATC CGCGGGGCTCCAGAGTCCCTACCACCAGGTCTGCGAGCTCATTGTGCTGGAGAACTC ACACAGCATCTCGGGGGTCCGGATCCTCCCCATCAACTACCTACGCCAGGTGCACCTC CTGGCCCAGACCTATGGGGCTGGGTCCACCTGGACAGGGGCCAGCTGATGAACGTG TGCTGGCCCTGCATGTCCCGCCACCTACATTGTGGAGCACTGTGACTCTGTGTCTTT CTGCCTCTCCAAGGGCTGGGCGCCCCAGTGGGGTCCCTGGCTGGGAGGCCCAAGGA TTCATTGAAGAAGCCTGGCGCTCCAAAAAGCCTGGGTGGAAGCATGCACCAGGTG GGATGCTGGCTGCGGTGGCCCTGCTGTGCCGTGGGGGTCCCCAGCAGGCCTGCAG CTTCTGATCCTGGGACTCAGGGTGGCCATTGAGTCACTAGGGCGGGAGCCAGGGCAC CAGCTGCTTCAGCAGACCCTACTCAGCGCAGAGTCCAGCCTGGGGAAGCCCTCGGTG ACGGCCAGGATTGTGATGCTCTGCAGAGTCTGGCAGAGCCTTAACCTGGGTCTTCC TCTGCTCACTGAGTCTCCAACAATCCCATGATGTCCCTACTTCACAATGAGGAACA GAGGCCAGAGAGGTCAATAG		
	ORF Start: ATG at 1	ORF Stop: TAG at 1063	
	SEQ ID NO: 140	354 aa	MW at 38439.8kD
NOV45a, CG96557-01 Protein Sequence	MLRRLARAAAVVGLRSDKVTRPGPAMRCVMAEAVVGGDDYGEDRVDELQEMAVELL LGRTLFPVSNMTANLISGDDCRCLGCVSTKKQHCSFSSRRCSVEGPPGAEGWSALP LPPRAAPLHSLAELERMITRGLQSPYHQVCELI CLENSHSISGVRI LPINYL RQVHL LAQTYGAWVHLDRQALMNVVLALHVPPTYIVEHCDVSFCLSKGLGAPVGS LAGRPKE FIEEAWRLQKALGGSMDHQVGM LAAVALLCRRGSPSRPAGFWILGLRVAIESLGRDPG QLLQQTLLSAESSLGKPSVDGQDCDALQSLAEPLTWVFHLLTESSNNPMMSLLHNEET EAQRGG		

- 5 Further analysis of the NOV45a protein yielded the following properties shown in Table 45B.

Table 45B. Protein Sequence Properties NOV45a	
PSort analysis:	0.8673 probability located in mitochondrial matrix space; 0.5542 probability located in mitochondrial inner membrane; 0.5542 probability located in mitochondrial intermembrane space; 0.5542 probability located in mitochondrial outer membrane
SignalP analysis:	Cleavage site between residues 17 and 18

A search of the NOV45a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 45C.

Table 45C. Geneseq Results for NOV45a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV45a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG42377	Arabidopsis thaliana protein fragment SEQ ID NO: 52844 - Arabidopsis thaliana, 356 aa. [EP1033405-A2, 06-SEP-2000]	12..260 4..248	98/252 (38%) 129/252 (50%)	4e-36
AAG42376	Arabidopsis thaliana protein fragment SEQ ID NO: 52843 - Arabidopsis thaliana, 358 aa. [EP1033405-A2, 06-SEP-2000]	12..260 6..250	98/252 (38%) 129/252 (50%)	4e-36
AAG42378	Arabidopsis thaliana protein fragment SEQ ID NO: 52845 - Arabidopsis thaliana, 339 aa. [EP1033405-A2, 06-SEP-2000]	26..260 1..231	88/238 (36%) 119/238 (49%)	5e-31
AAG10781	Arabidopsis thaliana protein fragment SEQ ID NO: 9239 - Arabidopsis thaliana, 332 aa. [EP1033405-A2, 06-SEP-2000]	26..260 1..231	94/241 (39%) 125/241 (51%)	5e-31
AAG73780	Human colon cancer antigen protein SEQ ID NO:4544 - Homo sapiens, 272 aa. [WO200122920-A2, 05-APR- 2001]	10..268 3..251	91/261 (34%) 135/261 (50%)	2e-30

In a BLAST search of public sequence databases, the NOV45a protein was found to
 5 have homology to the proteins shown in the BLASTP data in Table 45D.

Table 45D. Public BLASTP Results for NOV45a.				
Protein Accession Number	Protein/Organism/Length	NOV45a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
T00716	L-allo-threonine aldolase homolog F22O13.11 - Arabidopsis thaliana, 484 aa.	12..260 90..334	98/252 (38%) 129/252 (50%)	1e-35
AAL86346	HYPOTHETICAL 38.9 KDA PROTEIN - Arabidopsis thaliana (Mouse-ear cress), 358 aa.	12..260 6..250	98/252 (38%) 129/252 (50%)	1e-35
Q9FRS2	F22O13.11 - Arabidopsis thaliana (Mouse-ear cress), 352 aa.	12..260 6..250	98/252 (38%) 129/252 (50%)	1e-35
O07051	L-allo-threonine aldolase (EC 4.1.2.-) (L-allo-TA) (L-allo-threonine acetaldehyde-lyase) - Aeromonas jandaei, 338 aa.	12..259 4..241	100/249 (40%) 131/249 (52%)	3e-34
Q8ZGE4	L-ALLO-THREONINE ALDOLASE (EC 4.1.2.-) - Yersinia pestis, 339 aa.	11..268 2..249	93/260 (35%) 136/260 (51%)	1e-31

PFam analysis predicts that the NOV45a protein contains the domains shown in the Table 45E.

Table 45E. Domain Analysis of NOV45a			
Pfam Domain	NOV45a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 46.

- 5 The NOV46 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 46A.

Table 46A. NOV46 Sequence Analysis			
	SEQ ID NO: 141	945 bp	
NOV46a, CG96581-01 DNA Sequence	TCTGCCCAGTTAAAGTGCAGGCGTTAGGCTAAAGCTGGGACAAACCTGGAGCGGCTTC ATTGGCTTTCTCTACTTCTTTTGCATCCAGAGATGCATAAGCTTAAATCGTCTCAGAA GGACAAGGTCGCCCAGTTTATGGCGTGCACTCAGGCTGGCGAGAGAAGTCTCTATCTAC TGCTTAACGCAGAATGAGTGGAGACTAGACGAGGCCACGGACAGGACTTCTTCCAAAACC CAGACTCGCTCCACAGGGAGTCCATGCGGAACGCTGTGGACAAGAAGAAGCTGGAGCG GCTGTACGGCAGGTACAAAGATCCACAAGATGAAAACAAAATTGGAGTCGATGGGATT CAACAGTTTTGTGATGATCTGAGCCTGGATCCTGCCAGTATCAGTGTATTGGTCATAG CGTGGAAAGTTCAGGGCAGCAACTCAGTGTGAATTTAGCAGAAAGGAATTTCTAGATGG CATGACAGAACTTGGGTGTGACAGCATGGAGAAGCTAAAGGCTCTTCTGCCAAGACTG GAGCAGGAGCTGAAGGACACAGCCAAGTTTAAAGATTTTTATCAGTTTACCTTCACCT TCGCTAAGAACCAGGGCAGAAAGGTTTAGACTTAGAAATGGCTGTTGCGTATTGGAA ATTAGTGTTATCTGGAAGGTTTAAATTTTATAGATCTCTGGAACACATTCTTAATGGAA CATCAAAAAGATCAATTCCAAGGGACACCTGGAACCTCCTGCTGGACTTTGGAAACA TGATTGCGGATGATATGTCTAACTACGATGAAGAAGGAGCTTGGCCCGTTCTTATAGA TGATTTTGTAGAATATGCACGGCCAGTAGTCACAGGTGGAAAACGCAGCCTTTCTTAG GCAGCAAGTTAAGCAGGAGTAAGATTATGAAATGATTGTATCTGCAAGGAGATTGC AGTCAGTTCCTGGGTGC		
	ORF Start: ATG at 91	ORF Stop: TAG at 868	
	SEQ ID NO: 142	259 aa	MW at 30179.1kD
NOV46a, CG96581-01 Protein Sequence	MHKLKSSQKDKVRQFMACTQAGERTAICYCLTQNEWRLDEATDSFFQNPDSLHRESMRN AVDKKKLERLYGRYKDPQDENKIGVDGIQQFCDDLSLDPASISVLVIAWKFRATQCE FSRKEFLDGMTELGCDSMEKLKALLPRLEQELKDTAKFKDFYQPTFTFAKNPQGKGLD LEMAVAYWKLVLSGRFKFLDLWNTFLMEHHKRSIPRDTWNLLDFGNMIADMSNYDE EGAWPVLIDDFVEYARPVVTGGKRSLF		
	SEQ ID NO: 143	808 bp	
NOV46b, CG96581-02 DNA Sequence	GAGATGCATAAGCTTAAATCGTCTCAGAAGGACAAGGTCGCCAGTTTATGGCGTGCA CTCAGGCTGGCGAGAGAACTGCTATCTACTGCTTAACGCAGAAATGAGTGGAGACTAGA CGAGGCCACGGACAGCTTCTTCCAAAACCCAGACTCGCTCCACAGGGAGTCCATGCGG AACGCTGTGGACAAGAAGAAGCTGGAGCGGCTGTACGGCAGGTACAAAGATCCACAAG ATGAAAACAAAATTGGAGTCGATGGGATTCAACAGTTTGTGATGATCTGAGCCTGGG TCCTGCCAGTATCAGTGTATTGGTCATAGCGTGGAAGTTCAGGGCAGCAACTCAGTGT GAATTTAGCAGAAAGGAATTTCTAGATGGCATGACAGAACTTGGGTGTGACAGCATGG AGAAGCTAAAGGCTCTTCTGCCAAGACTGGAGCAGGAGCTGAAGGACACAGCCAAGTT TAAAGATTTTTATCAGTTTACCTTCACCTTCGCTAAGAACCAGGGCAGAAAGGTTTA GACTTAGAAATGGCTGTGCGTATTGGAAATTAGTGTTATCTGGAAGGTTTAAATTTT TAGATCTCTGGAACACATTCTTAATGGAACATCAAAAAGATCAATTCGAAGGGACAC CCGGAACCTCCTGCTGGACTTTGGAACATGATTGCGGATGATATGTCTAACTACGAT GAAGAAGGAGCTTGGCCCGTTCTTATAGATGATTTGTAGAATATGCACGGCCAGTAG TCACAGGTGGAAAAACGCAGCCTTTTCTAGGCAGCAGTTAAGCAGGAGTAAGATT		
	ORF Start: ATG at 4	ORF Stop: TAG at 781	
	SEQ ID NO: 144	259 aa	MW at 30149.1kD
NOV46b, CG96581-02 Protein Sequence	MHKLKSSQKDKVRQFMACTQAGERTAICYCLTQNEWRLDEATDSFFQNPDSLHRESMRN AVDKKKLERLYGRYKDPQDENKIGVDGIQQFCDDLSLDPASISVLVIAWKFRATQCE FSRKEFLDGMTELGCDSMEKLKALLPRLEQELKDTAKFKDFYQPTFTFAKNPQGKGLD LEMAVAYWKLVLSGRFKFLDLWNTFLMEHHKRSIPRDTNRNLLDFGNMIADMSNYDE EGAWPVLIDDFVEYARPVVTGGKRSLF		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 46B.

Table 46B. Comparison of NOV46a against NOV46b.		
Protein Sequence	NOV46a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV46b	1..259	258/259 (99%)
	1..259	258/259 (99%)

Further analysis of the NOV46a protein yielded the following properties shown in Table 46C.

Table 46C. Protein Sequence Properties NOV46a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

- 5 A search of the NOV46a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 46D.

Table 46D. Geneseq Results for NOV46a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV46a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM40345	Human polypeptide SEQ ID NO 3490 - Homo sapiens, 244 aa. [WO200153312-A1, 26- JUL-2001]	16..259 1..244	244/244 (100%) 244/244 (100%)	e-144
AAM42131	Human polypeptide SEQ ID NO 7062 - Homo sapiens, 224 aa. [WO200153312-A1, 26- JUL-2001]	2..204 22..224	200/203 (98%) 200/203 (98%)	e-115
AAV94962	Human secreted protein clone mt124_3 protein sequence SEQ ID NO:130 - Homo sapiens, 244 aa. [WO200009552-A1, 24-FEB-2000]	16..255 1..240	185/240 (77%) 218/240 (90%)	e-115
AAB93082	Human protein sequence SEQ ID NO:11917 - Homo sapiens, 186 aa. [EP1074617-A2, 07- FEB-2001]	1..173 1..173	173/173 (100%) 173/173 (100%)	1e-98
ABB63270	Drosophila melanogaster polypeptide SEQ ID NO 16602 - Drosophila melanogaster, 291 aa. [WO200171042-A2, 27- SEP-2001]	2..248 5..247	152/248 (61%) 192/248 (77%)	1e-87

In a BLAST search of public sequence databases, the NOV46a protein was found to have homology to the proteins shown in the BLASTP data in Table 46E.

Table 46E. Public BLASTP Results for NOV46a				
Protein Accession Number	Protein/Organism/Length	NOV46a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96GG9	RP42 HOMOLOG - Homo sapiens (Human), 259 aa.	1..255 1..255	201/255 (78%) 234/255 (90%)	e-123
Q9QZ73	RP42 - Mus musculus (Mouse), 259 aa.	1..255 1..255	201/255 (78%) 235/255 (91%)	e-123
Q9HCT3	RP42 PROTEIN - Homo sapiens (Human), 259 aa.	1..255 1..255	200/255 (78%) 234/255 (91%)	e-123
Q99NE7	PUTATIVE LEUCINE-ZIPPER PROTEIN - Mus musculus domesticus (western European house mouse), 259 aa.	1..255 1..255	200/255 (78%) 233/255 (90%)	e-122
AAL78673	LEUCINE ZIPPER PROTEIN - Homo sapiens (Human), 259 aa.	1..255 1..255	199/255 (78%) 232/255 (90%)	e-122

PFam analysis predicts that the NOV46a protein contains the domains shown in the Table 46F.

Table 46F. Domain Analysis of NOV46a			
Pfam Domain	NOV46a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 47.

- 5 The NOV47 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 47A.

Table 47A. NOV47 Sequence Analysis

	SEQ ID NO: 145	1072 bp	
NOV47a, CG96624-01 DNA Sequence	CCCTCCTTGTCTGTCTCCGGGATT CAGGCCTCCCTCCCTGACATGGAGAGTAACCTGT CTGGCCTGGTGCCTGCTGCCGGGCTGGTGCCTGCGCTGCCACCTGCTGTGACCTCTGGG GCTGACAGCTGCCTACACCACCTGTATGCCCTGCTCTTCTCCGCTCTATGCCAG CTCTGGCTGGTGCCTTCTGTATGGGCACAAGCGTCTCAGCTATCAGACGGTGTCTCTGG CCCTCTGTCTGTCTCTGGGCCGCCTTGCCTGACCACCTCTTCTCCTTCTACTTCCGAGA TACTCCCCGCGCAACCGCCTGGGGCCCTTGCCCTTCTGGCTTCTCTACTGCTGCCCC GTCTGCCTGCAGTTCTTACCTTGACGCTTATGAACCTTACTTTGCCAGGTAAGGC TCGCTGTCCGAGGGGCCTTTGTGGGGGCCTCGCTGCTCTTTCTGCTGGTGAACGTGCT GTGTGCTGTGCTCTCCCATCGGCGCCGGGCACAGCCCTGGGGCCTGCTGCTTGTCCGC GTCTGGTGAGCGACTCCCTGTTCTGTCATCTGCGCGCTGTCTCTTGTGCTGCTGCTCT GCCTCGTCGCCAGGCGGGCGCCCTCCACTAGCATCTACCTGGAGGCCAAGGGGACCAG TGTGTGCCAGGCGCGCGATGGGTGGCGCCATGGTCTGCTCTATGCCAGCCGGGCC TGCTACAACCTGACAGCACTGGCCTTGCGCCCCCAGAGCCGGCTGGACACCTTCGATT ACGACTGGTACAATGTGTCTGACAGGCGGACCTGGTGAATGACCTGGGGAACAAAGG CTACCTGGTATTGGCCCTCATCCTCTTCTGTGTGGGAGCTACTGCCACCACCTGCTG GTGGGCTTCTTCCGGGTGCACCGGCCCCACAGGACCTGTTTGCTCTCGGTCTTACT TCTTTGACGGGCTGGGCACCTGTGAAGATGAGGGCTGCTCCTGGGAGCACAGCCGGG TGAGAGCACCAGGTAGGAGCGGTGGCACTGCCTCAGTACCCCTGCCCTACCCGCCAC CCCGCTGGCTCCATCAAGCTATGGGGGA		
	ORF Start: ATG at 43	ORF Stop: TAG at 1000	
	SEQ ID NO: 146	319 aa	MW at 35534.2kD
NOV47a, CG96624-01 Protein Sequence	MESNLSGLVPAAGLVPAALPPAVTLGLTAAYTTLVALLFFSVYAQLWLVLVLYGHKRLSY QTVFLALCLLWALRTTLFSFYFRDTPRANRLGLPFWLLYCCPVLCQFFTLTLMNLY FAQVRLAVRGAFVGSLLFLLVNLCAVLSHRRRAQPPWALLVRVLVSDSLFVICALS LAACLCLVARRAPSTSIYLEAKGTSVCQAAMGGAMVLLYASRACYNLTALALAPQSR LDTFDYDWNVSDQADLVNDLGNKGYLVFGLILFVWELLPTLLVGFFRVHRPPQDLF ASRSYFFDRAGHCEDEGCSWEHSRGESTR		
	SEQ ID NO: 147	1200 bp	
NOV47b, CG96624-02 DNA Sequence	GATAAGGCCTCCCTCCCTGACATGGAGAGTAACCTGTCTGGCCTGGTGCCTGCTGCCG GGCTGGTGCCTGCGCTGCCACCTGCTGTGACCCTGGGGCTGACAGCTGCCTACACCAC CCTGTATGCCCTGCTCTTCTTCTCCGTCTATGCCAGCTCTGGCTGGTGCCTTCTGTAT GGGCACAAGCGTCTCAGCTATCAGACGGTGTCTCTGGCCCTCTGTCTGCTCTGGGCGG CCTTGGCTACCACTCTTCTCCTTCTACTTCCGAGATACTCCCCGCGCCAACCGCCT GGGGCCCTTGCCCTTCTGGCTTCTCTACTGCTGCCCGCTGCTGCTGAGTCTTCTCACC TTGACGCTTATGAACCTCTACTTTGCCAGGTAAGGCTCGCTGTCCGAGGGGCTTTG TGGGGGCTCGCTGCTCTTCTGCTGGTGAACGTGCTGTGTGCTGTGCTCTCCCATCG GCGCGGGCACAGCCCTGGGCCCTGCTGCTTGTCCGCGTCTCGGTGAGCGACTCCCTG TTCGTATCTGCGCGCTGTCTCTTGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTG CCTCCACTAGCATCTACCTGGAGGCCAAGGGGACCAAGTGTGTGCCAGGCGCGCGAT GGGTGGCGCATGCTGCTCTATGCCAGCCGGGCTGCTACAACCTGACAGCACTG GCCTTGGCCCCCAGAGCCGGCTGGACACCTTCGATTACGACTGGTACAATGTGTCTG ACCAGGTGGGCATACGCATGTCTGCCACCCAGGCGGACCTGGTGAATGACCTGGGGAA CAAAGGCTACCTGGTATTGGCCTCATCCTCTTCTGTGTGGGAGCTACTGCCACCACC CTGCTGGTGGGCTTCTTCCGGGTGCACCGGCCCCACAGGACCTGTTTGCTCTCGGT CCTACTTCTTTGACCGGGCTGGGCACTGTGAAGATGAGGGCTGCTCTGGGAGCACAG CCGGGCTGAGGACCAAGTATGTGGGCGAGTCTAGGCTCTGGGAGCTGGTATGGTGCC ATCGGGCGTGAGCCGGGCTGGTATGGGGGAGCCAGACGAAGACCTCCTCTGCTCT TCTCCAGGTGCCAGGACAGGCGGCCACCACCACTCTCTACTCCACCCACAGAC GTGATCCCCGTGCGCTGCTGGGAAGAATACCGAGGCCCTG		
	ORF Start: ATG at 22	ORF Stop: TGA at 1162	
	SEQ ID NO: 148	380 aa	MW at 41810.1kD
NOV47b, CG96624-02 Protein Sequence	MESNLSGLVPAAGLVPAALPPAVTLGLTAAYTTLVALLFFSVYAQLWLVLVLYGHKRLSY QTVFLALCLLWALRTTLFSFYFRDTPRANRLGLPFWLLYCCPVLCQFFTLTLMNLY FAQVRLAVRGAFVGSLLFLLVNLCAVLSHRRRAQPPWALLVRVLVSDSLFVICALS LAACLCLVARRAPSTSIYLEAKGTSVCQAAMGGAMVLLYASRACYNLTALALAPQSR LDTFDYDWNVSDQVGIRMSATQADLVNDLGNKGYLVFGLILFVWELLPTLLVGFFR VHRPPQDLFASRSYFFDRAGHCEDEGCSWEHSRGESTSMSGSLGSGSWYGAIGREPGW YGGSQTKTPLLFSQVPGPGHHHSLYSTPQT		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 47B.

Table 47B. Comparison of NOV47a against NOV47b.		
Protein Sequence	NOV47a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV47b	1..314	256/323 (79%)
	1..323	256/323 (79%)

Further analysis of the NOV47a protein yielded the following properties shown in Table 47C.

Table 47C. Protein Sequence Properties NOV47a	
PSort analysis:	0.6850 probability located in endoplasmic reticulum (membrane); 0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 44 and 45

- 5 A search of the NOV47a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 47D.

Table 47D. Geneseq Results for NOV47a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV47a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE09061	Human seven-transmembrane protein 19459 - Homo sapiens, 396 aa. [WO200159116-A2, 16-AUG-2001]	1..318 1..343	317/343 (92%) 317/343 (92%)	e-180
AAU04571	Human G-protein coupled receptor like protein, GPCR #8 - Homo sapiens, 1314 aa. [WO200153454-A2, 26-JUL-2001]	1..318 1..343	317/343 (92%) 317/343 (92%)	e-180
AAY69992	Human receptor-associated protein from Incyte clone 786873 - Homo sapiens, 346 aa. [WO200008155-A2, 17-FEB-2000]	1..318 1..293	267/343 (77%) 267/343 (77%)	e-142
AAU04572	Human G-protein coupled receptor model sequence #4 - Homo sapiens, 178 aa. [WO200153454-A2, 26-JUL-2001]	56..218 1..178	162/178 (91%) 162/178 (91%)	2e-85
AAY60207	Human endometrium tumour EST encoded protein 267 - Homo sapiens, 296 aa. [DE19817948-A1, 21-OCT-1999]	26..289 1..278	159/279 (56%) 197/279 (69%)	2e-85

In a BLAST search of public sequence databases, the NOV47a protein was found to have homology to the proteins shown in the BLASTP data in Table 47E.

Table 47E. Public BLASTP Results for NOV47a				
Protein Accession Number	Protein/Organism/Length	NOV47a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC69286	SEQUENCE 1 FROM PATENT WO0159116 - Homo sapiens (Human), 396 aa.	1..318 1..343	317/343 (92%) 317/343 (92%)	e-180
Q96N19	CDNA FLJ31532 FIS, CLONE NT2RJ2000597 - Homo sapiens (Human), 417 aa.	1..319 1..344	317/344 (92%) 317/344 (92%)	e-179
Q9NQC5	HYPOTHETICAL 23.3 KDA PROTEIN PRECURSOR - Homo sapiens (Human), 210 aa.	1..187 1..201	182/202 (90%) 182/202 (90%)	4e-96
Q9JHD9	PUTATIVE SEVEN PASS TRANSMEMBRANE PROTEIN - Mus musculus (Mouse), 385 aa.	14..310 18..339	180/323 (55%) 221/323 (67%)	2e-93
O60478	PUTATIVE SEVEN PASS TRANSMEMBRANE PROTEIN (SIMILAR TO TRANSMEMBRANE 7 SUPERFAMILY MEMBER 1) (UPREGULATED IN KIDNEY) - Homo sapiens (Human), 399 aa.	14..310 33..354	177/323 (54%) 220/323 (67%)	4e-93

PFam analysis predicts that the NOV47a protein contains the domains shown in the Table 47F.

Table 47F. Domain Analysis of NOV47a			
Pfam Domain	NOV47a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 48.

- 5 The NOV48 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 48A.

Table 48A. NOV48 Sequence Analysis			
	SEQ ID NO: 149	834 bp	
NOV48a, CG96747-01 DNA Sequence	ATGGCCCGGCTGGGCGCGCTGCTCCTGGCCGCGCCCTGGGTGCACTGCTCAGCTTCGCGCTCCTGGCCGCGCGGTGCGCAGCGACTACTGGTACATCCTGGAGGTGGCGGACGCGGCAATGGCAGCGCCTGGCCCGGGCGCGCAGAGCTGCTCTCTCGCACTCGGGGCTCTGGCGCATCTGCGGAAGACCTCCTGACATGCACCTGTGTGCACGTGTGCACACGCGAGGCCTATGTTTTCCAGGGCAGAACGGCTGCATCCCGCTGGTTCGACCTTTTGCCAGTGA GAGCCTGGACGTCTCCACCTCGGTGCAGCACCTCATCTTGCTGCACCGTGCAGTCATTGTGGTCCTGCCCTGAGCCTGGTCCTTCTCGTGTGTGGCTGGATCTGCGGCCTGCTCAGCTCCCTGGCCAGAGCGTGTCTCTGCTGCTTTTACCCTGCTGCTACTTCTGCTGGGGGTGTCTGACACTGGCGGGGGTTCAGCATCTACATCAGCTACTCGCACCTGGCCTTTGCGGAGACGGTGCAGCAGTATGGCCCGCAGCACATGCAGGGCGTCCGCGTCAGCTTCGGCTGGTCCATGGCCCTGGCCTGGGGCTCCTGTGCCTTGGAGGCATTACAGCGGAACCTCCTGCTCTCAGCTGCTGGACCCTCAGCCTGAGCCCCCAATCTGTGGTCATCTGAGTCCCAGCAGGGCAGGCTTGGTGGAGAAGAGGCTGATGAGAGGGCCCGAGAGCCCTCCGATTTCAGGGGTGGGGGGCAAGGAGCTGAGCGATCCAGATGTACCCCTCTGCCCCCTCCCTTGTCTCAAGCCTGCTAG		
	ORF Start: ATG at 1	ORF Stop: TAG at 832	
	SEQ ID NO: 150	277 aa	MW at 29226.6kD
NOV48a, CG96747-01 Protein Sequence	MARLGALLLAAALGALLSFALLAAVASDYWYILEVADAGNGSAWPGRAELSSHSGLWRICGRPPDMHLCARVHTRGLCFPGQNGCIPLVDPFASELDVSTSVQHLILLHRAVIVLPLSLVLLVCGWICGLLSSLAQSVSLLFTGCFYLLGGVLTLAGVSIYISYSHLAF AETVQQYGPQHMQGVRVSPGWSMALAWGSCALEAFSGTLLLSAANTLSLSPICGHLS PQQGRLGGEADERAREPLRFAGVGGKELSDPDVPLCPLPCSQAC		

Further analysis of the NOV48a protein yielded the following properties shown in Table 48B.

Table 48B. Protein Sequence Properties NOV48a	
PSort analysis:	0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.3700 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 29 and 30

A search of the NOV48a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 48C.

Table 48C. Geneseq Results for NOV48a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV48a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM23670	Human EST encoded protein SEQ ID NO: 1195 - Homo sapiens, 57 aa. [WO200154477-A2, 02-AUG-2001]	186..241 1..57	51/57 (89%) 53/57 (92%)	1e-22
AAB73979	Human stargazin-like protein CACNG8 - Homo sapiens, 327 aa. [WO200121791-A2, 29-MAR-2001]	5..168 7..161	49/172 (28%) 81/172 (46%)	2e-09
AAY84372	A human voltage-gated calcium channel designated CACNGLIKE3 - Homo sapiens, 327 aa. [WO200014224-A1, 16-MAR-2000]	5..168 7..161	49/172 (28%) 81/172 (46%)	2e-09
AAY70462	Human membrane channel protein-12 (MECHP-12) - Homo sapiens, 323 aa. [WO200012711-A2, 09-MAR-2000]	5..168 7..158	42/169 (24%) 79/169 (45%)	3e-09
ABB11805	Human voltage gated Ca channel subunit homologue, SEQ ID NO:2175 - Homo sapiens, 325 aa. [WO200157188-A2, 09-AUG-2001]	5..168 9..160	42/169 (24%) 78/169 (45%)	7e-09

In a BLAST search of public sequence databases, the NOV48a protein was found to have homology to the proteins shown in the BLASTP data in Table 48D.

Table 48D. Public BLASTP Results for NOV48a				
Protein Accession Number	Protein/Organism/Length	NOV48a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9D563	493051 IJ11RIK PROTEIN - Mus musculus (Mouse), 222 aa.	3..224 4..215	98/231 (42%) 133/231 (57%)	2e-38
Q9JJV4	Voltage-dependent calcium channel gamma-4 subunit (Neuronal voltage-gated calcium channel gamma-4 subunit) - Mus musculus (Mouse), 327 aa.	5..168 7..161	50/172 (29%) 81/172 (47%)	3e-09
Q98UH4	IPR328-LIKE PROTEIN - Gallus gallus (Chicken), 314 aa.	1..168 3..157	49/171 (28%) 78/171 (44%)	4e-09
Q8WXS5	VOLTAGE-DEPENDENT CALCIUM CHANNEL GAMMA-8 SUBUNIT - Homo sapiens (Human), 426 aa.	5..216 18..226	59/232 (25%) 98/232 (41%)	5e-09
CAC36506	SEQUENCE 1 FROM PATENT WO0121791 - Homo sapiens (Human), 327 aa.	5..168 7..161	49/172 (28%) 81/172 (46%)	6e-09

PFam analysis predicts that the NOV48a protein contains the domains shown in the Table 48E.

Table 48E. Domain Analysis of NOV48a			
Pfam Domain	NOV48a Match Region	Identities/ Similarities for the Matched Region	Expect Value
PMP22_Claudin	4..214	48/232 (21%) 139/232 (60%)	0.023

Example 49.

- 5 The NOV49 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 49A.

Table 49A. NOV49 Sequence Analysis			
	SEQ ID NO: 151	537 bp	
NOV49a, CG96789-01 DNA Sequence	CGCCTCCCGGCGCGGGTGGCCGAGGCGTAGCGCTGCGACCCCGCACCCCTCTGCGAA CATGGCGCTGCGAGTGGTGGCGAGCGTGCGGGCCCTGCTCTGCACCCCTGCGCGCGGTG CCGTACCCCGCCGCGCCCTGCCCGCCGAGGCCCTGGCAGCTGGGGGTGGGCGCCGTCC GTACGCTGCGCACTGGACCCGCTCTGCTCTCGGTGCGTAAATTTACAGAGAAACACGA ATGGGTAAACAACAGAAAATGGCATTGGAACAGTGGGAATCAGCAATTTTGACAGGAA GCGTTGGGAGATGTTGTTTATTGTAGTCTCCCTGAAGTTGGGACAAAATTGAACAAAC AAGGTTGGCTGATCAAGATGACACTGAGTAACCCCTTCAGAACTAGATGAACCTATGAG TGAAGAAGCATATGAGAAATACATAAAATCTATTGAGGAGTGAAAATGGAACCTCCTAA ATAAACTAGTATGAAATAACGCAAGCCAGCAGAGTTGTCTTAAATTAGTGGTGGATAG AAGACTTAGAATAGC		
	ORF Start: ATG at 60	ORF Stop: TGA at 447	
	SEQ ID NO: 152	129 aa	MW at 14212.3kD
NOV49a, CG96789-01 Protein Sequence	MALRVVRSVRLCLTLRAVPSPAAPCPPRPWQLGVGAVRTLRTGPALLSVRKFTKHE WVTTENGIGTVGISNFAQEALGDVVYCSLPEVGTKLNKQGWLIKMTLSNPSELDELM EEAYEKYIKSIEE		

Further analysis of the NOV49a protein yielded the following properties shown in Table 49B.

Table 49B. Protein Sequence Properties NOV49a	
PSort analysis:	0.9200 probability located in mitochondrial matrix space; 0.6000 probability located in mitochondrial inner membrane; 0.6000 probability located in mitochondrial intermembrane space; 0.6000 probability located in mitochondrial outer membrane
SignalP analysis:	Cleavage site between residues 19 and 20

- 5 A search of the NOV49a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 49C.

Table 49C. Geneseq Results for NOV49a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV49a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG11725	Novel human diagnostic protein #11716 - Homo sapiens, 185 aa. [WO200175067-A2, 11-OCT-2001]	1..129 10..185	124/176 (70%) 126/176 (71%)	6e-61
ABG11725	Novel human diagnostic protein #11716 - Homo sapiens, 185 aa. [WO200175067-A2, 11-OCT-2001]	1..129 10..185	124/176 (70%) 126/176 (71%)	6e-61
ABG18844	Novel human diagnostic protein #18835 - Homo sapiens, 184 aa. [WO200175067-A2, 11-OCT-2001]	1..129 9..184	103/176 (58%) 111/176 (62%)	2e-46
ABG18844	Novel human diagnostic protein #18835 - Homo sapiens, 184 aa. [WO200175067-A2, 11-OCT-2001]	1..129 9..184	103/176 (58%) 111/176 (62%)	2e-46
AAB58488	Lung cancer associated polypeptide sequence SEQ ID 826 - Homo sapiens, 102 aa. [WO200055180-A2, 21-SEP-2000]	7..60 1..54	52/54 (96%) 52/54 (96%)	1e-24

In a BLAST search of public sequence databases, the NOV49a protein was found to have homology to the proteins shown in the BLASTP data in Table 49D.

Table 49D. Public BLASTP Results for NOV49a				
Protein Accession Number	Protein/Organism/Length	NOV49a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P23434	Glycine cleavage system H protein, mitochondrial precursor - Homo sapiens (Human), 173 aa.	1..129 1..173	128/173 (73%) 128/173 (73%)	5e-63
Q9N121	Glycine cleavage system H protein, mitochondrial precursor - Oryctolagus cuniculus (Rabbit), 173 aa.	1..129 1..173	114/173 (65%) 119/173 (67%)	1e-54
P20821	Glycine cleavage system H protein, mitochondrial precursor - Bos taurus (Bovine), 173 aa.	1..129 1..173	113/173 (65%) 118/173 (67%)	1e-53
Q96GY5	GLYCINE CLEAVAGE SYSTEM PROTEIN H (AMINOMETHYL CARRIER) - Homo sapiens (Human), 98 aa.	1..97 1..97	96/97 (98%) 96/97 (98%)	2e-50
Q9QYU8	H PROTEIN - Rattus norvegicus (Rat), 169 aa.	1..126 1..168	102/170 (60%) 112/170 (65%)	3e-46

PFam analysis predicts that the NOV49a protein contains the domains shown in the Table 49E.

Table 49E. Domain Analysis of NOV49a			
Pfam Domain	NOV49a Match Region	Identities/ Similarities for the Matched Region	Expect Value
GCV_H	51..96	25/51 (49%) 46/51 (90%)	2.3e-22
GCV_H	97..127	14/35 (40%) 31/35 (89%)	2.4e-11

Example 50.

- 5 The NOV50 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 50A.

Table 50A. NOV50 Sequence Analysis			
	SEQ ID NO: 153	440 bp	
NOV50a, CG97253-01 DNA Sequence	CCTGGACTCAATCATGGCTTGTGGTCTGGTCGCCAGCAACCTGAATCTCAAACCTGGA GAGTGCCCTTCGAGTGCAGGCGCAGGTGGCTCCTGACGCTAAGAGCTTCGTGCTGAACCC TGGGCAAGACAGCAACAACCTGTGCCTGCACTTCAACCCCTCGCTTCAACGCCACCGG CGACGCCAACACCATCGTGTGCAACAGCAAGGACGGCGGGGCTGGGGGACCGAGCAG CGGGAGGCTGTCTTTCCCTTCCAGCCTGGAAGTGTGTGCAGAGGTGTGCATCACCTTCG ACCAGGCCAACCTGACGGTGACTTCAAGATCAAATGTGTGGCCTTTGACTGAAATCAG CCAGCCCATGGCCCCCAATAAAGGCAGCTGCCTCTGAAAAAAAAAAAAAAAAAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 14	ORF Stop: TGA at 383	
	SEQ ID NO: 154	123 aa	MW at 13177.8kD
NOV50a, CG97253-01 Protein Sequence	MACGLVASNLNLKPGCELRVRGEVAPDAKSFVLNLGKDSNNLCLHFNPRFNAHGDAANT IVCNSKDGGAWGTEQREAVFPFQPGSVAEVCITFDQANLTVTSRSNVWPLTEISQPM PNKGSCL		

Further analysis of the NOV50a protein yielded the following properties shown in Table 50B.

Table 50B. Protein Sequence Properties NOV50a	
PSort analysis:	0.6500 probability located in cytoplasm; 0.1586 probability located in lysosome (lumen); 0.1000 probability located in mitochondrial matrix space; 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV50a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 50C.

Table 50C. Geneseq Results for NOV50a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV50a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY80675	Human galectin-1 fusion protein, SEQ ID NO:9 - Homo sapiens, 142 aa. [WO200006724-A1, 10-FEB-2000]	1..99 8..106	99/99 (100%) 99/99 (100%)	1e-54
AAW55959	Human HL-60 lectin - Homo sapiens, 135 aa. [WO9808535-A2, 05-MAR-1998]	1..99 1..99	99/99 (100%) 99/99 (100%)	1e-54
AAR08153	Human lectin gene product - Homo sapiens, 135 aa. [JP02262597-A, 25-OCT-1990]	1..99 1..99	99/99 (100%) 99/99 (100%)	1e-54
AAR52745	Amino acid sequence of both HL-60 and placenta lectins - Homo sapiens, 135 aa. [WO9411497-A, 26-MAY-1994]	1..99 1..99	98/99 (98%) 99/99 (99%)	3e-54
AAP91964	HL-60 and placenta lectins - Homo sapiens, 135 aa. [EP337799-A, 18-OCT-1989]	1..99 1..99	98/99 (98%) 99/99 (99%)	3e-54

In a BLAST search of public sequence databases, the NOV50a protein was found to have homology to the proteins shown in the BLASTP data in Table 50D.

Table 50D. Public BLASTP Results for NOV50a				
Protein Accession Number	Protein/Organism/Length	NOV50a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9D0X0	18 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY, CLONE:1110058J24, FULL INSERT SEQUENCE - Mus musculus (Mouse), 152 aa.	1..123 1..149	103/149 (69%) 112/149 (75%)	2e-54
LNHUGB	galectin 1 [validated] - human, 135 aa.	1..99 1..99	99/99 (100%) 99/99 (100%)	3e-54
E981607	AMINO ACID SEQUENCE FOR GBP - vectors, 135 aa.	1..99 1..99	99/99 (100%) 99/99 (100%)	3e-54
CAA01896	BETA GALACTOSIDE BINDING PROTEIN - synthetic construct, 135 aa.	1..99 1..99	98/99 (98%) 98/99 (98%)	1e-53
P09382	Galectin-1 (Beta-galactoside-binding lectin L-14-I) (Lactose-binding lectin 1) (S-Lac lectin 1) (Galaptin) (14 kDa lectin) (HPL) (HBL) - Homo sapiens (Human), 134 aa.	2..99 1..98	98/98 (100%) 98/98 (100%)	1e-53

PFam analysis predicts that the NOV50a protein contains the domains shown in the Table 50E.

Table 50E. Domain Analysis of NOV50a			
Pfam Domain	NOV50a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Gal-bind_lectin	3..113	65/139 (47%) 100/139 (72%)	4.6e-45

Example 51.

- 5 The NOV51 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 51A.

Table 51A. NOV51 Sequence Analysis

	SEQ ID NO: 155	4297 bp
NOV51a, CG97400-01 DNA Sequence	GCAGGGCCCTGAAGACTGGCGATCCCGCGCCCGACTACCTTGGTCTCGCGGATATGC TCCATCGCAAGGAGACGTGAGCCGGCAGCCGGTCTCACAGGCGTACCATACGGCTACA GAAACAGGGCGGTGACAGCGACAGAGAGCAAGCGGGAACCTCCCTCCACCAGCCAGCGC CGCGCGTCTGCTGCTGCTCGAGTCCCCGGGAGGCGCGGGGTTGGGGAAGTGTCTTCT AGGAGACGGCGCTCACCGGCTGCACCTGCGCGGTGACGCCACCGGGGCCGCGAGACA GACCCGCGGCGCTGGCTGGTGGAGGGAGTTCCCGCTTGCTCTGTGCTGTACCGC CCTGTTTCTGTAGCCGTATGGTACGCCCTGTGAGACCGGCTGCCGGGTGACGCTCTCCTT GCGATGGAGCATATCCGGACGACCAAGGTCGAACAAGTAAATTAACCTGACCGGATTC GTACCAGCAACAAGTCATTAACAGGAACACTGTATCTTACGGCTACACATCTATTATT TATCGACTCTCATCAAAAGAAACCTGGATATTACACCACCATATTGCTCAGTAGAG AAACCTTGCTTGAAGTCTCTGGATGCCCCCTTGTGATACAGTGCAAGAACTTCAGAA CTGTGCTTTCATTTGTTCCAGAGAAAGAGATTGCCATGATATTACAACTCTTTGCT ACAACTGTCAAAACAAGCAAAATATGAAGATCTCTATGCATTTTCTTATAATCCCAA CAAAATGATTGAGAACGACTACAAGGCTGGCAGCTCATTGATCTCGCTGAGGAATATA AGAGGATGGGAGTGCCAACTCACACTGGCAGTTGTCTGATGCCAACCGGGACTACAA GATTTGTGAAACTTACCCAGAGAACTTTATGTTCCCGGATAGCAAGCAAAACCAATA ATTGTTGGTAGTTCCAAAGTCCGGAGCAAGGGAAGATTCCAGTTCTTTCTACTATC ATCAAGATAAGGAGGCTGCCATTGTGCGATGTAGTCAGCCACTCTCTGGATTTCAGTGC CAGGTGCCCTGGAGGATGAACATTGTCTTCAAGCCATTAGTAAAGCCAACTCAGTCAAT CGCTATATGTACGTGATGGATACAGGCCAAACTGAATGCAATGGCCACAGAGCAG CTGGAAGAGGTTATGAAATGAAGACAACCTATTCCAATATTAGATTTTCAGTTTGTGG AATTGAAAATATTTCATGTGATGAGGTCCAGCCTTCAGAAATTATGGAAGTCAATGGC ACTAAAGGGCTTTCTGTCAATGATTTCTACTCCGTTTGGAGAGCTCGGGATGGCTTC GCCATATCAAAGCTGTTATGGATGCTGCAGTCTTCTTGGCCAAAGCAATAACAGTTGA AAATGCAAGTGTGTTGGTGCATTGTTCCGATGTTGGGATAGGACTTCCAGGTTTGT TCCTGGGTTCTCTTTTATGGATTCTACTACAGGACAATCAAAGGATTCATGGTTT TAATAGAAAAGGATTGGATCTCTTTTGGACATAAATTTTTCAGAGAGGTGTGGCCAGTT GGATGGTGACCCAAAGGAAGTCTCACCAGTGTCTTACTCAGTTCTTGAATGTGTGTGG CATTGACCGAACAGTTTCCACAAGCCTTTGAATTCAGTGAAGCAATTTCTTCTTCAGA TCCATGAGCATATTTCATTTCATGCCAGTTTGGAAACTTCTTGGAAATTTGCCCAAGGA AAGAGAAGAGCTCAAGTTGAAGGAGAAGACTTATCCCTGTGGCCATTCTTTTGGAA GACCAAAAGAAAGTACTTAAATCCTCTCTACAGTTCCGAATCTCACAGATTTACAGTTT TGGAGCCAAATACAGTATCTTTCAATTTAAGTTTGGAGGAACATGTACCATCAGTT TGATCGAACACTGCATCTAGGCAGTCTGTATTTAATATAATTATGAATATGAATGAG CAAAATAAACAAATAGAGAAAGATATTAAAGACCTAGAATCTAAAATTAACCAACGCA AAAATAAGCAACAGATGGCATCTCCACCAAGGAATTGTACATTTCAGTTTCATCCTGA ATCACCTAACCTCAAAACTTCCCTGTGTTTAAAGAGCAGACTCTGCTACCCGTAAT GATGCTCTTGAAGTATAGAGGCGAGCAGCCCGGAGATAATCGTTATAGTGAATATG CAGAAGAGTTTCTAAATCAGAACCTGCTGTGGTGCAGTTAGAGTATGGTGTGGCAG AATGACTTGTAGACTCATAGAGTTTCTTCTGCAATGATTGCAGTACAAGAAAAGGAT TATTGTGAGGATGGTCTGTAAGCATAACCAAAAGGAATTGTCTAATAACAAATTTAG GGTAAACAGTAGGCTAATAGTTGAAGGAAGGATAATACTACCCTTGTGAGAGAAAT ATGTCAATTTAATTGCATTTCCAGCAAGGAATGACATTCAGTTCTGTGAAGAAATGAGT GGTATTTGATGTATTTACTCAAAACAAATTTGCACTGTACACTGATGAATTTGACGTT TATGATTTATGTTAATTGAGCAAAACATAAATAACCTTCTTAAAGTACAATTTAACTT CAAGAAAACAAAATTTGACAACATAGTTTCTTAAATAATGATATGGCATGTACTTTCA ATTATGTAGCTTTGTAACATGAATATTTACATATTTTGCCCTTTTAGTGATATTTAAT GTTAAAGTGCCATGAAAAATATTTCTAAGAAAGCCTTAAATCCCAGTGGATCTTTTA CCCTTAAGTTTACAGCTACAACAAGATTTTGTGTTTGTGTTTCTTCTGGTCAGCC TTGTTTGTGTTGTAAGAATGTGCTCTCATTACTGCTGGGGGTGCATGCTACATAC TTCTATATAAACACTTGTAGAAGTACACTGTTACGTTTACCTGCCCCACTTTTGTA TTCAAAATTAATGAACTGAAGGTTTATTCTGATCATAATTTGTTTGTGCTACATTT GATAATTTATTTTACAGCTTAGAATATTGATTCTTGAATACGTATAAGCACATTT GACTGTCTTTTATATATGGATTACTGCATTCCATTGATTCTTATTTGTGGTTGGCTT TATTTCTTTCACAAGTGTGTAAGTTTAAAGAGCTAAAGCTCTAAACCTGTTCTGAGAA ACAATGAATAGTACATGTATATGTATATTTTAAACTGCCTTATTGCTCAATGAGTTG CTGTTCCCTGGAGTATCTAATTTTCAGTTTACTCAAAACTTTTGGGTAAATAAAAAG GGAAGTACAGATAGTTTAGGTTGTGATGGTGCATGAATTTGAAGTCATTCTATG TGGAGCATATTTTCTTCTTGTAAATATAGAAAAAAGAGAGATTCTAGTACAAAG TTAAGTTTAAACAAAGCAACATAAATCTGGGAAAGATTTCAATTTGCCATGTTATA TTTACTGTTTATTCTGTGTACTAGTACATATCTTAAATTACCAAAACCAAGAAACA AAACATAAAAAACCCAAACTATCACTTGGAAATTAGCAATATACCAACCTGGCTTTTA AAATTGAAAATTTAAATAACATGGTGGCATGAAGTACAATTCGAGTATTAGGCAATTT GCATAGTGTCTCATGCTACTTTCTGTTACACCTCTATTATATTAGTTTGAATATA AACATCTTTTTCAGACCAAAAAAATTTTATTGTATGAGAGCTTATCTTATCCTGTTT ATTTTTCATGCTTTCTGTAATACATTATGTAAATTTAAATAATTTCTTTTAA CAGCAACAGAAATGCATATAAATATAGTATGTGATTAAACCAATCTGCTTCCATAT	

	TTAAGCACTGGGAATGGAACTTAATCTCTGTGACTACAAAGGGAAGTTTTGTGCCT TGGTGTTCAGTCAGTATTGTGGTTTTAGAACTCTCTGTGGCTGACTTCTGTATTCT AACTCTGGTTTTATTATCTTAATCTTCAACTTCAGACATACATTGTGTGTTTATATGC TCACAGGTGGACTGAGAAATCAGTTACATCTTAAGTGACCTACAGGGTATATGTTGGC AAAAGCAGACTGTGTATATGCTTATAAAGTGAATTTATGTTTCAGTGTGTTTGGAG TGTATAGCATGTAAATTATTTATATATGATTTAAAGGTAATTAATGTTTCACATTT ACTTTGAATGTTTTCTTCTGGATAAAACAAATGGAATCAAAAAAAAAAAAAAAAAA AAATT		
	ORF Start: ATG at 410		ORF Stop: TAG at 2273
	SEQ ID NO: 156	621 aa	MW at 71922.1kD
NOV51a, CG97400-01 Protein Sequence	MEHIRTTKVEQVKLLDRFSTSNKSLTGTLYLTATHLLFIDSHQKETWILHHHIASVEK LALTTSGCPLVIOCKNFRTVHFIVPRERDCHDIYNSLLQLSKQAKYEDLYAFSPYNPKQ NDSERLQGWQLIDLAEYKRMGVPNSHWQLSDANRDYKICETYPRELYVPRIASKPII VGSSKFRSKGRFPVLSYHQDKEAAICRCSQPLSGFSARCLEDEHLLQAIKANPVR YMYVMDTRPKLNAMANRAAGKGYENEDNYSNIRFQFVGIEINIHVMRSSLQKLEVN KGLSVNDFYSGLESSGWLRIKAVMDAAVFLAKAITVENASVLVHCSGDGWDRTSQVCS LGSLLDSYRTIKGFMVLIKDWISFGHKFSERCGQLDGDGPKEVSPVFTQPLECVWH LTEQFPQAFEFSEAFLLQIHEHIHSCQFGNFGNCPKEREELKLEKTYSLWPFLLD QKKYLNPLYSSSHRFTVLEPNTVSFNFKFWRNMYHQFDRTLHPRQSVFNIIMNMNEQ NKQLEKDIKDLESKIKQRKNKQTDGILTKELLHSHVHSPNLTSLCFKEQTLPLVND ALRTIEGSSPADNRYSEYAEFFSKSEPAVVSLEYGVARMTC		

Further analysis of the NOV51a protein yielded the following properties shown in Table 51B.

Table 51B. Protein Sequence Properties NOV51a	
PSort analysis:	0.3600 probability located in mitochondrial matrix space; 0.1723 probability located in microbody (peroxisome); 0.1000 probability located in lysosome (lumen); 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV51a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 51C.

Table 51C. Geneseq Results for NOV51a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV51a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB73229	Human phosphatase MTMR7_h - Homo sapiens, 629 aa. [WO200112819-A2, 22-FEB-2001]	8..534 82..613	338/533 (63%) 430/533 (80%)	0.0
ABB66818	Drosophila melanogaster polypeptide SEQ ID NO 27246 - Drosophila melanogaster, 676 aa. [WO200171042-A2, 27-SEP-2001]	14..542 1..525	282/530 (53%) 386/530 (72%)	e-176
ABB59964	Drosophila melanogaster polypeptide SEQ ID NO 6684 - Drosophila melanogaster, 676 aa. [WO200171042-A2, 27-SEP-2001]	14..542 1..525	282/530 (53%) 386/530 (72%)	e-176
ABB65465	Drosophila melanogaster polypeptide SEQ ID NO 23187 - Drosophila melanogaster, 639 aa. [WO200171042-A2, 27-SEP-2001]	52..542 2..488	260/491 (52%) 355/491 (71%)	e-163
AAW36451	Human tyrosine phosphatase related protein - Homo sapiens, 621 aa. [WO9735015-A2, 25-SEP-1997]	22..510 74..567	197/501 (39%) 297/501 (58%)	4e-99

In a BLAST search of public sequence databases, the NOV51a protein was found to have homology to the proteins shown in the BLASTP data in Table 51D.

Table 51D. Public BLASTP Results for NOV51a				
Protein Accession Number	Protein/Organism/Length	NOV51a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96P80	MYOTUBULARIN RELATED PROTEIN 6 - Homo sapiens (Human), 621 aa.	1..621 1..621	620/621 (99%) 620/621 (99%)	0.0
Q8VE11	HYPOTHETICAL 70.9 KDA PROTEIN - Mus musculus (Mouse), 617 aa.	1..621 1..617	558/621 (89%) 588/621 (93%)	0.0
Q9Y217	Myotubularin related protein 6 (EC 3.1.3.48) - Homo sapiens (Human), 465 aa (fragment).	157..621 1..465	464/465 (99%) 464/465 (99%)	0.0
Q96EF0	SIMILAR TO HYPOTHETICAL PROTEIN FLJ20126 - Homo sapiens (Human), 704 aa.	1..540 1..541	332/543 (61%) 430/543 (79%)	0.0
Q9W1Q6	BCDNA:GH04637 PROTEIN - Drosophila melanogaster (Fruit fly), 676 aa.	14..542 1..525	282/530 (53%) 386/530 (72%)	e-175

PFam analysis predicts that the NOV51a protein contains the domains shown in the Table 51E.

Table 51E. Domain Analysis of NOV51a			
Pfam Domain	NOV51a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 52.

- 5 The NOV52 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 52A.

Table 52A. NOV52 Sequence Analysis			
	SEQ ID NO: 157	807 bp	
NOV52a, CG97462-01 DNA Sequence	ATGGCTGCCAAAGTGTGAGTTCATCAGCAAGTTGGCCTGGCCTTAGCTGTTGCAG GAGGCCTGAATGAACTCTGCCTTAATGTGGATGCTGGGCACAGAGCTGTGCATCTTTGA CCTATTCCGTGGAGTACAGGACATTGTGGTAGGGGAAAGGACTCACTTCTCATTCCCA TGGGTACAGAAACCAATTATCTTTGACTGCCCTTCTCGACCAGTAAATGTGCCAGCCA TCACTGGTAGCAAAGATTACAGAATGTCAACATCACACTGCTCATCCTCTTCTGGCC TGTCACTAGCCAGTTTCTTGCATCTTACCAGCATCAGAGAGGACTATGATGAGCAG GTGCTGCCATCCGTACGACCAAGATCCTCAAGTCCGTGGTGGCTAGCTTTGATGCTG GAGAACTAATCACCAGAGAGAGCTGGTCTCCAGGGAGCAATTACTTACAGAGCGAGC AGCCACCTTTGGGCTCATCCTGGATGACGTGTCTCTTGACACATCTGACCTTCGGGAAG GAGTTCACAGAAGCGGTGGAAGCCAAACAGGTAGCTCAGCAGGAAGCAAGGGCCAGAT TTGTGGTGGAAAAGGCTGAGCAGCAGAAAAGGTGGCCATCATCTCTGCTGAGGGCTA CTCCAAGGCAGCTGAGCTGATTGCCAACTCACTGGCCACCGCAAGGGACCGCCTGATG GAGCTCTGCAAGCTGGAAGCTGCGGAGGACATCGCGTACCAGCTCTCACGCTCTCGGA ACATCACCTATCCGCGGCTGGGCAGTCCGTGCTCCTCCAGCTGCCCCAGTGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 805	
	SEQ ID NO: 158	268 aa	MW at 29706.9kD
NOV52a, CG97462-01 Protein Sequence	MAAKVFEFISKFGLALAVAGGLNELCLNVDAGHRAVIFDLFRGVQDIVVGERTHFLIF WVQKPIIFDCPSRPRNPVPAITGSKDLQNVNITLLILFWPVTSSQFPCIFTSIREYDEQ VLPSVTTKILKSVVASFDAGELITQRELVSREQLLTERAATFGLILDVSLTHLTFGK EFTEAVEAKQVAQQEARARFVVEKAEQQKKVAII SAEGYSKAELIANSLATARDRLM ELCKLEAAEDIAYQLSRSRNITYPPAGQSVLLQLPQ		

Further analysis of the NOV52a protein yielded the following properties shown in Table 52B.

Table 52B. Protein Sequence Properties NOV52a	
PSort analysis:	0.5500 probability located in endoplasmic reticulum (membrane); 0.2632 probability located in lysosome (lumen); 0.1000 probability located in endoplasmic reticulum (lumen); 0.1000 probability located in outside
SignalP analysis:	Cleavage site between residues 22 and 23

A search of the NOV52a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 52C.

Table 52C. Geneseq Results for NOV52a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV52a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG73845	Human colon cancer antigen protein SEQ ID NO:4609 - Homo sapiens, 279 aa. [WO200122920-A2, 05-APR-2001]	1..268 8..279	237/272 (87%) 243/272 (89%)	e-120
AAB43874	Human cancer associated protein sequence SEQ ID NO:1319 - Homo sapiens, 279 aa. [WO200055350-A1, 21-SEP-2000]	1..268 8..279	237/272 (87%) 243/272 (89%)	e-120
AAW54352	Heat shock 27 kD protein and prohibitin (admixture) - Homo sapiens, 471 aa. [WO9810291-A1, 12-MAR-1998]	1..268 200..471	237/272 (87%) 243/272 (89%)	e-120
AAR42215	Human prohibitin - Homo sapiens, 272 aa. [JP05271294-A, 19-OCT-1993]	1..268 1..272	237/272 (87%) 243/272 (89%)	e-120
AAR13466	Prohibitin - Rattus rattus, 272 aa. [USN7612674-N, 16-JUL-1991]	1..268 1..272	236/272 (86%) 243/272 (88%)	e-120

In a BLAST search of public sequence databases, the NOV52a protein was found to have homology to the proteins shown in the BLASTP data in Table 52D.

Table 52D. Public BLASTP Results for NOV52a				
Protein Accession Number	Protein/Organism/Length	NOV52a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P35232	Prohibitin - Homo sapiens (Human), 272 aa.	1..268 1..272	237/272 (87%) 243/272 (89%)	e-120
P24142	Prohibitin (B-cell receptor associated protein 32) (BAP 32) - Mus musculus (Mouse), and, 272 aa.	1..268 1..272	236/272 (86%) 243/272 (88%)	e-120
Q9VIZ4	LETHAL (2) 37CC PROTEIN - Drosophila melanogaster (Fruit fly), 276 aa.	1..267 1..271	178/271 (65%) 212/271 (77%)	4e-91
Q9BKU4	HYPOTHETICAL 30.0 KDA PROTEIN - Caenorhabditis elegans, 275 aa.	2..266 5..273	157/269 (58%) 203/269 (75%)	2e-78
O01720	Prohibitin-Like Molecule IC-PRO-I	2..267	155/270 (57%)	1e-76

	TC-PRO-1 - <i>Toxocara canis</i> , 274 aa.	4..273	201/270 (74%)	
--	--	--------	---------------	--

PFam analysis predicts that the NOV52a protein contains the domains shown in the Table 52E.

Table 52E. Domain Analysis of NOV52a			
Pfam Domain	NOV52a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Band_7	12..203	49/206 (24%) 169/206 (82%)	4.6e-66

Example 53.

The NOV53 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 53A.

Table 53A. NOV53 Sequence Analysis			
	SEQ ID NO: 159	1527 bp	
NOV53a, CG97472-01 DNA Sequence	ATTGAGGCCATGGGGACAAGAAGGTCACCCCATCTCTGATCTTCGCCATCACAAATTGCTACAATCGGCTCTTTCCAATTTGGCTACAACACTGGGGTCATCAATGCTCCTGAGATGATCATAAGGGAATTTATCAACACTTTGAAGGACAAGGCAAATACCCCTCCCTCTGAGATGTTTGTGCTTGTCCCTTTGGTGCTTGTCTATGGCCATATTCTCCATTGGTTGTATGATTAGCTCCTTTTCTGTTGGACTGTTTGTCAACAGCTTTGACAGGCGTAATTCAATGCTTACTGTCAACCTGTGGCTGCCACTGGTGGCTGCCTTATGGGACTGTGTAAAGTAGCTGAGTTGGTTGAAATACTGATCCTGGCCTACTTGGTTATTGGCCTCTTCTGCAGACTCTGCACAGGTTTTGTGCCACGTACATCGGAGACATCTCACCTATTGGCCTGCAAGGTGTTCTTTGGCACTCTCAACCAGCTGGGCATTGTTGTTGGAATTCTGGTGGCCAGATCTTTGATCTGGAATTCATCTCTGGGGTCTGAAGACCTATGGCCTGTGCTATTAGGCTTTCCCATCTCTCTGCTATGCTACAAAGTGACGCCCTTCTTTTGGCCCTGAAAGTCCCAGATTTCTTGCTCATTAAACAGAAAAGAAGAGGAGAATGCTAAGGAGATCCTCCAGTGGTTGTGGGGCACCCAGGATGTATCCCAAGACATCCAGGAGATGAAAGATGAGAGTGCAAGGATGGCACAAAGAAAGCAAGTCACTGTGCTGGAGCTCTTTAGAGTATCCAGCTACCGACAGCCCATCATCATTTCCATCATGCTCCAGCTCTCTCAGCAGCTCTCTGTAATCAATACTGTGTCTATTACTCAACAGGAATCTTTAAGGGTGCAAGGTGTTCAAGAGCCCATCTGTGTCAATTTGGTGTGGGTGTGGTTAATACTATCTTCACTATAGTTTCCCTATTTCTAGTGGAAAGGGCAGTAAGAAGGACTCTACATATGATAGGCCTTGGAGGGATGGCTTTTGTTCATCCTCATGTCTGTTTGTATTGTTATTGAAGGATGAGTGAATGGGATAAGCTTTGTCTGTATTGGGGCTATCTTGGTCTTTGTGGTCTTCTTTGAAATTGGGCCAGACCACATTCCCTGGTTTATTGTGGCTGAACTCTTCAGCCAGGGCCCTGGCCAGCTGTGATGGCAGTGGCCGGCTGCTCCACTGGACCTCCAACCTCCTAGTCAGATTGCTTTTCCCTTCTGCTGCTTACTATTTAGGAGCCTACGTTTATTATCTTCACCGACTTCCTCATTACCTTCTTGATCTTTACCTTCTTCAAAGTCCCTGAGATGTGTTACAGGACTTTCAAGGATATCACA AAGGCCTTTGAGGGGCAGGCACACGATGCAAAATAGATCTGGAAGGGACTGTGTATGAGATGAACAGTATCCAGCCTGCTAAGGAGACCACCACCAATGTCTAAATCATGCCTCC TTCAACCTCCCTCCAGCA		
	ORF Start: ATG at 10	ORF Stop: TAA at 1495	
	SEQ ID NO: 160	495 aa	MW at 54631.8kD
NOV53a, CG97472-01 Protein Sequence	MGTKKVTPSLIFAITIATIGSFQGYNTGVINAPEMI IREFINTLKDKNTPPSEMLCLSLWCLSMIAFISIGMISSFVGLFVNSFDRRNSMLTVNLLAATGGCLMGLCKVAEVLVEILILAYLVIGLFCRLCTGFVPTYIGDISPIGLQGVFTLNQLGIIVGILVAQIFDLEFILGSEDLWPVLLGFPISPAMLSAALPFCPEPRFLLINRKEEENAKEILQWLWGTQDVSQDIQEMKDESARMAQEKQVTVLELFRVSSYRQPIIISIMLQLSQLSVINTVFYY		

STGIFKAGVQEPICVTIGVGVNTIFTIVSLFLVERAVRRTLHMIGLGGMAFCSILM SVCLLLKDECNISFVCIGAILVFVVFVEIGPDHIPWFIVAELEFSQGPAPMAVAGC STWTSNLFVRLFPSPAAYLGAIFYFIIFTDFLITFLIFTFKVPENCYRTFKDITKAF EGQAHDANRSGRDCVMEMNSIQPAKETTTNV
--

Further analysis of the NOV53a protein yielded the following properties shown in Table 53B.

Table 53B. Protein Sequence Properties NOV53a	
PSort analysis:	0.6850 probability located in endoplasmic reticulum (membrane); 0.6400 probability located in plasma membrane; 0.4600 probability located in Golgi body; 0.1000 probability located in endoplasmic reticulum (lumen)
SignalP analysis:	Cleavage site between residues 22 and 23

A search of the NOV53a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 53C.

Table 53C. Geneseq Results for NOV53a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV53a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAY27289	Glucose transporter protein GLUT3 - Homo sapiens, 494 aa. [US5942398-A, 24-AUG-1999]	1..491 1..492	375/492 (76%) 420/492 (85%)	0.0
AAR11360	Glucose Transporter Protein from CHO cells - Cricetulus sp, 492 aa. [WO9103554-A, 21-MAR-1991]	3..474 5..478	276/474 (58%) 349/474 (73%)	e-154
ABB57244	Mouse ischaemic condition related protein sequence SEQ ID NO:652 - Mus musculus, 492 aa. [WO200188188-A2, 22-NOV-2001]	3..474 5..478	276/474 (58%) 347/474 (72%)	e-154
AAW17835	Human glucose transporter GLUT-1 - Homo sapiens, 492 aa. [WO9715668-A2, 01-MAY-1997]	3..474 5..478	275/474 (58%) 347/474 (73%)	e-153
AAB30522	Amino acid sequence of a consensus GLUT polypeptide - Synthetic, 493 aa. [US6136547-A, 24-OCT-2000]	6..487 10..490	272/482 (56%) 341/482 (70%)	e-150

In a BLAST search of public sequence databases, the NOV53a protein was found to have homology to the proteins shown in the BLASTP data in Table 53D.

Table 53D. Public BLASTP Results for NOV53a				
Protein Accession Number	Protein/Organism/Length	NOV53a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P11169	Solute carrier family 2, facilitated glucose transporter, member 3 (Glucose transporter type 3, brain) - Homo sapiens (Human), 496 aa.	1..495 1..496	418/496 (84%) 444/496 (89%)	0.0
AAL89710	GLUCOSE TRANSPORTER 14 LONG FORM - Homo sapiens (Human), 520 aa.	3..495 27..520	408/494 (82%) 436/494 (87%)	0.0
AAL89709	GLUCOSE TRANSPORTER 14 SHORT FORM - Homo sapiens (Human), 497 aa.	4..495 5..497	407/493 (82%) 436/493 (87%)	0.0
P47842	Solute carrier family 2, facilitated glucose transporter, member 3 (Glucose transporter type 3, brain) - Canis familiaris (Dog), 495 aa.	1..493 1..494	390/494 (78%) 431/494 (86%)	0.0
P47843	Solute carrier family 2, facilitated glucose transporter, member 3 (Glucose transporter type 3, brain) - Ovis aries (Sheep), 494 aa.	1..491 1..492	375/492 (76%) 420/492 (85%)	0.0

PFam analysis predicts that the NOV53a protein contains the domains shown in the Table 53E.

Table 53E. Domain Analysis of NOV53a			
Pfam Domain	NOV53a Match Region	Identities/ Similarities for the Matched Region	Expect Value
sugar_tr	12..464	173/489 (35%) 383/489 (78%)	3.5e-143

Example 54.

- 5 The NOV54 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 54A.

Table 54A. NOV54 Sequence Analysis

	SEQ ID NO: 161	2049 bp	
NOV54a, CG97528-01 DNA Sequence	TACACATTTTCCCTCATCTCTCTCACTGACACACACAAAACCTTTTGGCCATCCTGGGG AATTCTGTGACTCTATCTCTAATTGCAAAAGACTGGAAATCTAATTCATTCCAAC CTTCCCTTGACAGACAAGAGAACAAATGCCCTGGACATGGCTCCAGAGATCCACATGAC GGCCCAATGTGCCTCATTGAGAACACTAATGGGGAAGTGGTGGCGAATCCAGAAGCTC TGAAATCCTGTCTGCCATTACACAGCCTGTGGTGGTGGTGGCAATTTGTGGGCTCTA CCGCACAGGAAAATCCTACCTGATGAACAAGCTAGCTGGGAAGAATAAGGGCTTCTCT CTGGGCTCCACAGTGAAATCTCACACCAAGGAATCTGGATGTGGTGTGTGCCTCAC CCAAAAGCCAGAACACACCTTAGTCCTGCTTGACACTGAGGGCTGGGAGATGTAA GAAGGGTGACAACCAGAATGACTCCTGGATCTTCACCCTGGCCGCTCCTCTGAGCAGC ACTCTCGTGTAATAGCATGGGAACCATCAACCAGCAGGCTATGGACCAACTGCAGT ATGTGACAGAGCTGACACATCGAATCCGATCAAAATCCTCACCTGATGAGAATGAGAA TGAGGATTGAGCTGACTTTGTGAGCTTCTTCCAGATTTTGTGTGGACACTGAGAGAT TTCTCCCTGGACTTGAAGCAGATGGACAACCCCTCACACCAGATGAGTACCTGGAGT ATTCCCTGAAGCTAACGCAAGGTACCAGTCAAAAAGATAAAAATTTAATCTGTCCCGG ACTCTGTATCCGGAAGTTCTTCCCAAAGAAAAATGTTTGTCTTCGATCTGCCATT CACCGCAGGAAGCTTGCCAGCTTGAGAAACTACAAGATGAAGAGCTGGACCTGAAAT TTGTGCAACAAGTAGCAGACTTCTGTTCTCATCTTTAGCAATTTCCAAAACATAAAAC TCTTTCAGGAGGCATCAAGGTCAATGGGCCCTCGTCTAGAGAGCCTAGTGCTGACCTAT ATCAATGCTATCAGCAGAGGGGATCTGCCCTGCATGGAGAACGAGCTCTGGCCTTGG CCAGATAGAGAACTCAGCCGAGTGCAAAAGGCTATTGCCCACTATGACCAGCAGAT GGGCCAGAAGGTGCAGCTGCCCGCAGAAACCTCCAGGAGCTGCTGGACCTGCACAGG GTTAGTGAGAGGGAGGCCACTGAAGTCTATATGAAGAACTCTTTCAGGATGTGGACC ATCTGTTTCAAAGAAATTAAGGCCAGCTAGACAAAAGCGGGATGACTTTTGTAA ACAGAATCAAGAAGCATCATCAGATCGTTGCTCAGCTTTACTTCAGGTCAATTTTCAGT CCTCTAGAAGAAGAAGTGAAGGCGGAATTTATTGAAACAGGGGGCTATTGTCTCT TTATTGAGAAGCTACAAGACCTGGAGAAAAAGTACTATGAGGAACCAAGGAAGGGTCC TAAGGCTGAAGAGATTCTGCAGACATACTTGAATCCAGGAGTCTGTGACCGATGCA ATTCTACAGACAGACAGATTCTCACAGAAAAGGAAAAGGAGATTGAAGGTGTGGAAT GTGTAAGGCTGAATCTGCACAGGCTTACGCAAAAATGGTGGAGGAATGCAAAATAA GTATCAGCAGATGATGGAAGAGAAAGAGAAGAGTTATCAAGAACATGTGAAACAATTG ACTGAGAAGATGGAGAGGGAGAGGGCCAGTTGCTGGAAGAGCAAGAGAAGACCCTCA CTAGTAACTTCAGGAACAGGCCCGAGTACTAAAGGAGAGATGCCAAGGTGAAAGTAC CCAACCTTCAAATGAGATACAAAAGCTACAGAAGACCCTGAAAAAAAACCAAGAGA TATATGTCGCATAAGACTAAAGATCTAAACAACAGAGCTTTTCTGTATCTCTAACCCAA GGCATAACTGAAACAATTTTAGAATTTGGAACAAGTGTCACTATATTGATAATAATT AGATCTTGCATCATAACAC		
	ORF Start: ATG at 151	ORF Stop: TAA at 1939	
	SEQ ID NO: 162	596 aa	MW at 68176.6kD
NOV54a, CG97528-01 Protein Sequence	MAPEIHMTGPMCLIENTNGELVANPEALKILSAITQPVVVVAIVGLYRTGKSYLMNKL AGKNKGFSLGSTVKSHTKGIWMWCVPHPKPEHTLVLLDTEGLGDVKKGDNDQNSWIF TLAVLLSSTLVNSMGTINQQAMDQLQYVTELTTHRIRSKSSPDENENEDSADVVSFFP DFVWTLRDFSLDLEADGQPLTPDEYLEYSLKLTQGTSGQDKNFNLPRLCIRKFFPKKK CFVFDLPIHRRKLAQLEKLQDEELDPEFVQVQVADFCSYIFSNSTKTLSSGGIKVNGPR LESILVLTYYINAIISRGDLP MENAVLALAQIENSAAVQKAIHYDQMGQKVQLPAETL QELLDLHRVSEEREATEVYMKNSFKDVDHLFQKKLKAQLDKKRD DFCQKQNEASSDRCS ALLQVIFSPLEEEVKAGIYSKPGGYCLFIQKLQDLEKKYEEPRKGPKEEILQTYLK SKESVTDAILQTDQILTEKEKEIEGVECVKAESAQASAKMVEEMQIKYQOMMEEKEKS YQEHVKQLTEKMERERAQLLEBQKTLTSLKQEQARVLKERCCQGESTQLQNEIQKLQK TLKKKTKRYMSHKLKI		

Further analysis of the NOV54a protein yielded the following properties shown in Table 54B.

Table 54B. Protein Sequence Properties NOV54a	
PSort analysis:	0.8000 probability located in nucleus; 0.6000 probability located in endoplasmic reticulum (membrane); 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV54a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 54C.

Table 54C. Geneseq Results for NOV54a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV54a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM41334	Human polypeptide SEQ ID NO 6265 - Homo sapiens, 605 aa. [WO200153312-A1, 26-JUL-2001]	1..596 11..605	591/596 (99%) 592/596 (99%)	0.0
AAM41333	Human polypeptide SEQ ID NO 6264 - Homo sapiens, 605 aa. [WO200153312-A1, 26-JUL-2001]	1..596 11..605	591/596 (99%) 592/596 (99%)	0.0
ABB12327	Human guanylate binding protein homologue, SEQ ID NO:2697 - Homo sapiens, 605 aa. [WO200157188-A2, 09-AUG-2001]	1..596 11..605	591/596 (99%) 592/596 (99%)	0.0
AAU25444	Human mddt protein from clone LG:444850.9:2000MAY19 - Homo sapiens, 570 aa. [WO200162922-A2, 30-AUG-2001]	1..554 11..563	547/554 (98%) 549/554 (98%)	0.0
AAM39547	Human polypeptide SEQ ID NO 2692 - Homo sapiens, 592 aa. [WO200153312-A1, 26-JUL-2001]	1..585 1..586	514/587 (87%) 550/587 (93%)	0.0

In a BLAST search of public sequence databases, the NOV54a protein was found to have homology to the proteins shown in the BLASTP data in Table 54D.

Table 54D. Public BLASTP Results for NOV54a

Protein Accession Number	Protein/Organism/Length	NOV54a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P32455	Interferon-induced guanylate-binding protein 1 (Guanine nucleotide- binding protein 1) - Homo sapiens (Human), 592 aa.	1..585 1..586	508/587 (86%) 547/587 (92%)	0.0
P32456	Interferon-induced guanylate-binding protein 2 (Guanine nucleotide- binding protein 2) - Homo sapiens (Human), 591 aa.	1..583 1..582	441/583 (75%) 508/583 (86%)	0.0
Q9NV33	CDNA FLJ10961 FIS, CLONE PLACE1000588, HIGHLY SIMILAR TO INTERFERON-INDUCED GUANYLATE-BINDING PROTEIN 1 - Homo sapiens (Human), 447 aa (fragment).	144..582 10..447	434/439 (98%) 435/439 (98%)	0.0
Q01514	Interferon-induced guanylate-binding protein 1 (Guanine nucleotide- binding protein 1) (Interferon-gamma inducible protein MAG-1) - Mus musculus (Mouse), 589 aa.	1..579 1..578	392/579 (67%) 481/579 (82%)	0.0
Q96PP8	GUANYLATE BINDING PROTEIN 5 - Homo sapiens (Human), 586 aa.	1..582 1..577	394/582 (67%) 469/582 (79%)	0.0

PFam analysis predicts that the NOV54a protein contains the domains shown in the Table 54E.

Table 54E. Domain Analysis of NOV54a

Pfam Domain	NOV54a Match Region	Identities/ Similarities for the Matched Region	Expect Value
GBP	6..280	196/284 (69%) 256/284 (90%)	1.6e-192
GBP_C	282..579	169/305 (55%) 255/305 (84%)	2.5e-144

Example 55.

The NOV55 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 55A.

Table 55A. NOV55 Sequence Analysis			
	SEQ ID NO: 163	1184 bp	
NOV55a, CG97629-01 DNA Sequence	TATCTAGAGATGCTTTAAATCCTGAGGGGTTGGAGGACTTAGCTTACGTGCTAAGAT GCTAGAGTCTGGTTTCTCACTCTCATGGCTGGGGACGTAAATCTCAAGCGAACCGCT ATGCAAACTGGACTTCTTCGGACAGGGACATTTTGCTACAGCCTATAAGGCCAGAGA TAAGAACACCAACCAATCGTCACCATTAAAGAAATCAAACCTTGGACACAGGTCAGAA GCTAAAGGTGGTATAAATAGAACAGCCTTAAGGGAGATAAAGCTCTTACAGAACTAA GTCATCCGAACACAATTGGTCTCCTTGACACATTGGACATAAGTCTAACATTAGCCT TGTCTTGGATGTTATGGAACTGACCTAGAGGTTATAATAAAGGATAACAGTCTTGTG CTGACACCATCCACATTCAAGCCTACATGTTGATGGCTTTTCAGGGCTTAGAATATT TGCATCAGCATTGGATTCTACACAGGGATCTGAAACCAAGGACTTCTTGTAGATAA GAATGGAGTTCGAAACTGGCAGACTTTGTCTGGTCAAATCATTGGGAGCCCCAGC AGGACTTACACACATCAAGTTGTGACCAGATGGTATTGGGGCCCTGTGGAGTTATTGT TTGGAAGTAGGATGTATGATGTGGGAATGGACATGTGGGCTGTTTGTGTTATTAGC AGAATTGCTTCTAAGGGTTCCTTTTTTGCCTGGAGAGTCAGACCTTGATCAGCTAACA AGAATATTTGAACTTTGGGTACACCAACTGAAGAACAGTGGCCTGACATGTGTAGCC TTCCAGATTCTGTGACATTTAAGAGTTTCCCTGGGGTCCCCTACAACACATCTTCAC TGCAGCTGGGGACGACTTGTCTGGATCTCATCCAAGGCTTGTCTTCTTTAAACCGTGT ACTAGGATCACAGCCTCACAGGCATTGAAGACCAAGGACTTCAGAACTGTCCAGGGC CAACACCTGGATGCCAGCTTCCAAGACCAAACTGTTTCAGTGAAGCATTAAAGGAACC GTCCAATCCAACCATGGCAACAAACAGAAAAAGACAGAGGCCTTAGAACAAGGAATA TTGCCCAAGAAGCTATGTTTTTAGTTCCAGTGAACATGGACAGTTCCACTGCTGAAA GAAATGATCCAAAGGCAAATAATA		
	ORF Start: ATG at 83	ORF Stop: TAG at 1124	
	SEQ ID NO: 164	347 aa	MW at 39095.0kD
NOV55a, CG97629-01 Protein Sequence	MAGDVKSQANRYAKLDFFGQGHFATAYKARDKNTNQIVTIKKIKLGRSEAKGGGINRT ALREIKLLQKLSHPNTIGLLDTFGHKSNISLVLDVMEVDLEVIKDNLSVLTPSHIQ YMLMAFQGLEYLHQHWILHRDLKPKDFLLDKNGVLKLADFVLVKSFGSPSRITYTHQV TRWYWAPVELLFGTRMYDVGMDMWAVCCLLAEILLRVPFLPGESDLDQLTRIFETLGT PTEEQWPDMSLPDSVTFKSPFPGVPLQHIPTAAGDLDLQGLFFNPCTRITASQA LKTKDFRNCPGPTPGCQLPRPNCVSVEALKEPSNPTMATKQKTEALEQGILPKKLCF		

Further analysis of the NOV55a protein yielded the following properties shown in Table 55B.

Table 55B. Protein Sequence Properties NOV55a	
PSort analysis:	0.7000 probability located in plasma membrane; 0.5582 probability located in microbody (peroxisome); 0.2000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV55a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 55C.

Table 55C. Geneseq Results for NOV55a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV55a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB85505	Human protein kinase SGK149 - Homo sapiens, 332 aa. [WO200155356-A2, 02-AUG-2001]	23..347 11..332	268/325 (82%) 285/325 (87%)	e-153
AAB53379	Human colon cancer antigen protein sequence SEQ ID NO:919 - Homo sapiens, 260 aa. [WO200055351-A1, 21-SEP-2000]	94..347 8..260	214/254 (84%) 227/254 (89%)	e-124
AAG75087	Human colon cancer antigen protein SEQ ID NO:5851 - Homo sapiens, 260 aa. [WO200122920-A2, 05-APR-2001]	94..347 8..260	214/254 (84%) 227/254 (89%)	e-124
ABB59809	Drosophila melanogaster polypeptide SEQ ID NO 6219 - Drosophila melanogaster, 353 aa. [WO200171042- A2, 27-SEP-2001]	1..339 1..342	201/344 (58%) 252/344 (72%)	e-113
AAG14009	Arabidopsis thaliana protein fragment SEQ ID NO: 13712 - Arabidopsis thaliana, 348 aa. [EP1033405-A2, 06- SEP-2000]	10..319 11..316	151/310 (48%) 209/310 (66%)	2e-85

In a BLAST search of public sequence databases, the NOV55a protein was found to have homology to the proteins shown in the BLASTP data in Table 55D.

Table 55D. Public BLASTP Results for NOV55a

Protein Accession Number	Protein/Organism/Length	NOV55a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q03147	Cell division protein kinase 7 (EC 2.7.1.-) (CDK-activating kinase) (CAK) (TFIIH basal transcription factor complex kinase subunit) (39 kDa protein kinase) (P39 Mo15) (Protein-tyrosine kinase MPK-7) (CR4 protein kinase) (CRK4) - Mus musculus (Mouse), 346 aa.	1..347 1..346	298/347 (85%) 313/347 (89%)	e-173
P50613	Cell division protein kinase 7 (EC 2.7.1.-) (CDK-activating kinase) (CAK) (TFIIH basal transcription factor complex kinase subunit) (39 kDa protein kinase) (P39 Mo15) (STK1) (CAK1) - Homo sapiens (Human), 346 aa.	1..347 1..346	292/347 (84%) 308/347 (88%)	e-169
S34652	cell division cycle-2-related protein kinase (EC 2.7.1.-) CRK4 - mouse, 346 aa.	1..347 1..346	292/347 (84%) 309/347 (88%)	e-169
P51952	Cell division protein kinase 7 (EC 2.7.1.-) (CDK-activating kinase) (CAK) (TFIIH basal transcription factor complex kinase subunit) (39 protein kinase) (P39 Mo15) - Rattus norvegicus (Rat), 329 aa (fragment).	9..338 1..329	281/330 (85%) 295/330 (89%)	e-163
S51085	Cdk-activating kinase Cdk7 - rat, 312 aa (fragment).	9..320 1..311	269/312 (86%) 281/312 (89%)	e-156

PFam analysis predicts that the NOV55a protein contains the domains shown in the Table 55E.

Table 55E. Domain Analysis of NOV55a

Pfam Domain	NOV55a Match Region	Identities/ Similarities for the Matched Region	Expect Value
pkinase	12..296	99/301 (33%) 217/301 (72%)	1.4e-61

Example 56.

The NOV56 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 56A.

Table 56A. NOV56 Sequence Analysis			
	SEQ ID NO: 165	1821 bp	
NOV56a, CG97648-01 DNA Sequence	CGGGAAGGGAAAGCAGCCAGCAGCCCTCCAGCCCTCTGTGCTTTCCCTGGGAGTGGC CCCCGTGCTCAGCCATGGTGGACATGGGGGCCCTGGACAACCTGATCGCCAACACCGC CTACCTGCAGGCCCGGAAGCCCTCGGACTGCGACAGCAAAGAGCTGCAGCGCGCGCGG CGTAGCCTGGCCCTGCCCGGGCTGCAGGGCTGCGCGGAGCTCCGCCAGAAGCTGTCCC TGAACCTCCACAGCCTGTGTGAGCAGCAGCCCATCGGTGCGCGCCTCTTCCGTGACTT CCTAGCCACAGTGCCACGTTCCGCAAGGCGGCAACCTTCTAGAGGACGTGCAGAAC TGGGAGCTGGCCGAGGAGGGACCCACCAAGACAGCGCGCTGCAGGGGCTGGTGGCCA CTTGTGCGAGTGCCCTGCCCGGGGAACCCGCAACCTTCTCAGCCAGGCCCTGGC CACCAAGTGCCAAAGCAGCCACCACTGAGGAAGAGCGAGTGGCTGCAGTGACGCTGGCC AAGGCTGAGGGCCATGGCTTCTTGCAAGAGCAGCCCTTTAAGGATTCGTGACCAGCG CCTTCTACGACAAGTTTCTGCAGTGGAACCTTTCGAGATGCAACCAAGTGTGAGACAA GTACTTCACTGAGTTCAGAGTGTGGGAAAGGTGGTTTGGGGAGGTATGTGCGCTC CAGGTGAAAAACACTGGGAAGATGTATGCCTGTAAGAACTGGACAAGAAGCGGCTGA AGAAGAAAGGTGGCGAGAAGATGGCTCTCTTGGAAAAGGAAATCTTGGAGAAGGTTCAG CAGCCCTTTTCATTGTCTCTCTGGCCTATGCCTTTGAGAGCAAGACCCATCTCTGCCTT GTCTATGAGCCTGATGAATGGGGAGACCTCAAGTTCACATCTACAACGTGGGCACGC GTGGCCTGGACATGAGCCGGGTGATCTTTTACTCGGCCAGATAGCCTGTGGGATGCT GCACCTCCATGAACTCGGCATCGTCTATCGGGACATGAAGCCTGAGAATGTGCTTCTG GATGACCTCGGCAACTGCAGGTTATCTGACCTGGGGCTGGCGTGGAGATGAAGGGTG GCAAGCCCATCACCAGAGGGCTGGAACCAATGGTTACATGGCTCCTGAGATCCTAAT GGAAAAGGTAAGTTATTCTCTATCCTGTGGACTGGTTTGGCATGGGATGAGCATTAT GAAATGGTTGCTGGACGAACACCATTCAAAGATTACAAGGAAAAGGTCAGTAAAGAGG ATCTGAAGCAAAGAACTCTGCAAGACGAGGTCAAATTCAGCATGATAACTTCACAGA GGAAGCAAAGATATTTGCAGGCTCTTCTTGGCTAAGAAACAGAGCAACGCTTAGGA AGCAGAGAAAAGTCTGATGATCCAGGAAACATCATTCTTTAAACGATCAACTTTC CTCGCTTGAAGCTGGCCTAATTGAACCCCATTTGTGCCAGACCTTCAGTGGTTTA TGCCAAAGACATCGCTGAAATTGATGATTTCTCTGAGGTTTCGGGGGGTGAATTTGAT GACAAAGATAAGCAGTCTTCAAAAACCTTTGCGACAGGTGCTGTTCCTATAGCATGGC AGGAAGAAATTATAGAAACGGGACTGTTTGAGGAACGTAATGACCCCAACAGACCTAC GGGTTGTGAGGAGGGTAATTCATCCAAGTCTGGCGTGTGTTTATTGTAAATTGCT CTCTTTACCAGACAGGCAGCAGGAGTCTCGGCTGACATAATCCTCGAATGTTCCACAC GTGGAATCTGTGGAATGAGGGC		
	ORF Start: ATG at 73	ORF Stop: TAA at 1732	
	SEQ ID NO: 166	553 aa	MW at 62210.8kD
NOV56a, CG97648-01 Protein Sequence	MVDMGALDNLIAANTAYLQARKPSDCDSKELQRRRSLALPGLQGCAELRQKLSLNFHS LCEQQPIGRRLFRDFLATVPTFRKAATFLEDVQNWELAEAGPTKDSALQGLVATCASA PAPGNPQPFLSQAVATKCAATTEEERVAAVTLAKAEAMAFLEQPPKDFVTSAFYDK FLQWKLFEQPVSDKYFTEFRVLGKGGFGEVCAVQVNTGKMYACKLDDKRLKKKGG EKMALLEKEILEKVSSPFIVSLAYAFESKTHLCLVMSLMNGGDLKFHYNVGTRGLDM SRVIFYSAQIACGMLHLHELGIYRDMKPENVLLDDLGNCRSLDLGLAVEMKGGKPI QRAGTNGYMAPEILMEKVSYSYPVDWFAMGCSIYEMVAGRTPFKDYKEKVSKECLKQR TLQDEVKFQHDNFTTEAKDICRLFLAKKPEQRLGSREKSDPRKHFFKTIINFPRL GLIEPPFVDPSPVYAKDIAEIDDFSEVRGVEFDDKDKQFFKNFATGAVPIAWQBEII ETGLFEELNDPNRPTGCEEGNSKSGVCLLL		

Further analysis of the NOV56a protein yielded the following properties shown in
5 Table 56B.

Table 56B. Protein Sequence Properties NOV56a	
PSort analysis:	0.9685 probability located in nucleus; 0.1101 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV56a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 56C.

Table 56C. Geneseq Results for NOV56a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV56a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE16273	Human kinase PKIN-19 protein - Homo sapiens, 553 aa. [WO200196547-A2, 20-DEC-2001]	1..553 1..553	553/553 (100%) 553/553 (100%)	0.0
AAG77815	Human G-protein coupled receptor kinase 1 protein - Homo sapiens, 553 aa. [WO200168869-A2, 20-SEP-2001]	1..553 1..553	553/553 (100%) 553/553 (100%)	0.0
AAU03502	Human protein kinase #2 - Homo sapiens, 553 aa. [WO200138503-A2, 31-MAY-2001]	1..553 1..553	551/553 (99%) 551/553 (99%)	0.0
AAG77816	Human G-protein coupled receptor kinase 2 protein - Homo sapiens, 353 aa. [WO200168869-A2, 20-SEP-2001]	1..350 1..350	350/350 (100%) 350/350 (100%)	0.0
AAY24423	GRK4 polymorphism GRK4-alpha protein sequence - Homo sapiens, 578 aa. [WO9935279-A1, 15-JUL-1999]	7..531 3..526	253/528 (47%) 348/528 (64%)	e-140

In a BLAST search of public sequence databases, the NOV56a protein was found to have homology to the proteins shown in the BLASTP data in Table 56D.

Table 56D. Public BLASTP Results for NOV56a				
Protein Accession Number	Protein/Organism/Length	NOV56a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8WTQ7	G-PROTEIN-COUPLED RECEPTOR KINASE 7 (G PROTEIN-COUPLED RECEPTOR KINASE 7) - Homo sapiens (Human), 553 aa.	1..553 1..553	553/553 (100%) 553/553 (100%)	0.0
Q8WMV0	RETINA G PROTEIN-COUPLED RECEPTOR KINASE 7 - Bos taurus (Bovine), 552 aa.	1..553 1..552	475/553 (85%) 508/553 (90%)	0.0
Q8WP15	G PROTEIN-COUPLED RECEPTOR KINASE 7 - Sus scrofa (Pig), 553 aa.	1..553 1..553	470/553 (84%) 509/553 (91%)	0.0
Q9Z2G7	G PROTEIN-COUPLED RECEPTOR KINASE GRK7 - Spermophilus tridecemlineatus (Thirteen-lined ground squirrel), 548 aa.	2..553 1..548	468/552 (84%) 507/552 (91%)	0.0
O73658	OLGRK-C - Oryzias latipes (Medaka fish), 557 aa.	1..553 1..557	330/563 (58%) 416/563 (73%)	0.0

PFam analysis predicts that the NOV56a protein contains the domains shown in the Table 56E.

Table 56E. Domain Analysis of NOV56a			
Pfam Domain	NOV56a Match Region	Identities/ Similarities for the Matched Region	Expect Value
RGS	55..78	13/24 (54%) 22/24 (92%)	5.6e-07
pkinase	191..454	94/294 (32%) 202/294 (69%)	5.2e-70
pkinase_C	455..471	8/17 (47%) 13/17 (76%)	0.3

Example 57.

- 5 The NOV57 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 57A.

Table 57A. NOV57 Sequence Analysis			
	SEQ ID NO: 167	1722 bp	
NOV57a, CG97658-01 DNA Sequence	GCATGGTGGCACCATCTACCACAGGTCTCCCCAGCCGATAGTGATGGAGGCCTGGA CGAGGCTGAAGGGCTCCAGGACTCACAGAGAGAGATGCCGCCACCCCTCTCCCTCG CCGCCCTCAGATCCAGCTCAGAAGCCACCACCTCGAGGCGCTGGGAGCCACTCCCTCA CTGTACAGGAGCAGCCTGTGCCTGTTTCGCTGCCTCACAGTTCCTGCTTGCTGTGGGGT GCTCTGGTTCAGCGGTTATGGCCACATGTGGTCACAGAAGCCGACAAACCTCGTCTCC TCTTTGCTGACGCTCCTGAAACAGCTGGAACCCACGTCCTGGCTTGACTCTGGGACGT GGGGAGTCCCGGTCTGCTGCTGGTCTTTCTGTCCGTGGGCGCTGGTCTCGTTACCAC CCTGGTGTGGCACCCTCTGAGGACACCCCCAGAGCCACCCACCCACTGCCCTGAG GACAGGCGCCAGTCAGTGAGCGCCAGCCCTCCTTCACTACTCAGAGTGGATGGAGG AGAAGATCGAGGATGACTTCCTGGACCTCGACCCGGTGCCCGAGACTCCTGTGTTTGA TTGTGTGATGGACATCAAGCCTGAGGCTGACCCCACTCACTCACCCTCAAGTCCATG GGTCTGCAGGAGAGGAGGGGTTCCAATGTCTCCCTGACCCTGGACATGTGCACTCCGG GCTGCAACGAGGAGGGGCTTTGGCTATCTCATGTCCCCACGTGAGGAGTCCGCCCGCA GTACCTGCTCAGCGCCTCCCGTGTCTCCAAGCAGAAGAGCTTCATGAAAAGGCCCTG GACCCCTTCTGCTGCAGGCGGAATTCTTTGAAATCCCCATGAACTTTGTGGTACCGA AAGAGTACGACATCCCTGGGAGGTGTGCGGAAGAACCAGGTACAAAACCATACTTCCCA CCCTCAGCAGCAGAGTGTGTCTGACCTCACCAGACCCCTGACGACCCCTCTGAGTTCCTA ATCAATGCCAACTACATCCGGGCTATGGTGGGGAGGAGAAGGTGTACATCGCCACTC AGGGACCCATCGTCAGCAGGTGCGCGACTTCTGGCGCATGGTGTGGCAGGAGCACAT GCCCATCATTTGTATGATACCAACATCGAGGAGATGAACGAGAAATGCACCGAGTAT TGGCCCGAGGAGCAGGTGGCGTACGACGGTGTGAGATCACTGTGCAGAAAGTCATTTC ACACGGAGGATTACCGGCTGCGACTCATCTCCCTCAAGAGTGGGACTGAGGAGCGAGG CCTGAAGCATTACTGGTTCACATCCTGGCCCGACCAAGACCCAGACCGGGCCCCC CCACTCCTGCACCTGGTGCAGGAGGTGGAGGAGGCAGCCAGCAGGAGGGTCCCCACT GTGCCCCATCATCGTCCACTGCAGTGCAGGAGTGGGAGGACCGGCTGCTTCATTGC CACCAGCATCTGTGCCAGCAGTGCAGGAGGGTGTAGTGGACATCTGAAGACC ACGTGCCAGCTCCGTACAGACAGGGGCGGCATGATCCAGCACTGCGAGCAGTACCAGT TTGTGCAACCACGTATGAGCCTCTACGAAAAGCAGCTGTCCCACCACTCCCAGAATG ACTGCGGTTCTCTACAAGTTCCTGGGCACTGCCAGCCTGAGTCTCGGCCCTCAC CAGGGCCCTGCCTCGGGTCGGAATTCGATATCAAGCTTAT		
	ORF Start: ATG at 3	ORF Stop: TGA at 1623	
	SEQ ID NO: 168	540 aa	MW at 60799.6kD
NOV57a, CG97658-01 Protein Sequence	MVAPSTTGLPQPIVMEALDEAELQDSQREMPPPPPPSPSPDPAQKPPPRGAGSHSL VRSSLCLFAASQFLLACVGLWFSGYGHMWSQNAITNLVSSLLTLKQLEPTSWLDSGT GVPGLLLVFLSVGLVLTTLVWHLRLTPPEPPTPLPPEDRRQSVSRQPSFTYSEWME KIEDDFLDLPVPETPVFDCVMDIKPEADPTSLTVKSMGLQERRGSNVSLTLDMCTPC CNEEGFGYLMSPRESAREYLLSASRVLQAEELHEKALDPFLLQAEFFEIFPMNFVVP EYDIPGRCKRNRYKTILPNPHSRVCLTSPDDPLSSYINANYIRGYGGEKVYIATQ GPIVSTVADFWRMVWQHEPTPIVMITNIEEMNEKCTEYWPEEQVAYDGEVITQKVIH TEDYRLRLISLKSGETEERGLKHYWFTSWPDQKTPDRAPPLHLVREVEEAAQEGPH APIIVHCSAGIGRTGCFIATSICCQQLRQEGVVDILKTTCLRQDRGGMQHCQEQYQ VHVMSLYEKQLSHQSPE		

Further analysis of the NOV57a protein yielded the following properties shown in Table 57B.

Table 57B. Protein Sequence Properties NOV57a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.1000 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV57a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 57C.

Table 57C. Geneseq Results for NOV57a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV57a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAG67630	Amino acid sequence of a human protein - Homo sapiens, 537 aa. [WO200109316-A1, 08-FEB-2001]	8..540 5..537	532/533 (99%) 532/533 (99%)	0.0
AAG67451	Amino acid sequence of a human polypeptide - Homo sapiens, 537 aa. [WO200109345-A1, 08-FEB-2001]	8..540 5..537	532/533 (99%) 532/533 (99%)	0.0
AAM39481	Human polypeptide SEQ ID NO 2626 - Homo sapiens, 565 aa. [WO200153312-A1, 26-JUL-2001]	8..540 33..565	527/533 (98%) 529/533 (98%)	0.0
AAM41267	Human polypeptide SEQ ID NO 6198 - Homo sapiens, 600 aa. [WO200153312-A1, 26-JUL-2001]	8..540 68..600	526/533 (98%) 528/533 (98%)	0.0
AAE14457	Human protein phosphatase-7 - Homo sapiens, 541 aa. [WO200196546-A2, 20-DEC-2001]	8..540 9..541	526/533 (98%) 528/533 (98%)	0.0

In a BLAST search of public sequence databases, the NOV57a protein was found to
 5 have homology to the proteins shown in the BLASTP data in Table 57D.

Table 57D. Public BLASTP Results for NOV57a

Protein Accession Number	Protein/Organism/Length	NOV57a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
P54829	Protein-tyrosine phosphatase, non-receptor type 5 (EC 3.1.3.48) (Protein-tyrosine phosphatase striatum-enriched) (STEP) (Neural-specific protein-tyrosine phosphatase) - Homo sapiens (Human), 537 aa (fragment).	8..540 5..537	532/533 (99%) 532/533 (99%)	0.0
P54830	Protein-tyrosine phosphatase, non-receptor type 5 (EC 3.1.3.48) (Protein-tyrosine phosphatase striatum-enriched) (STEP) (Neural-specific protein-tyrosine phosphatase) - Mus musculus (Mouse), 541 aa.	8..540 9..541	474/533 (88%) 495/533 (91%)	0.0
P35234	Protein-tyrosine phosphatase, non-receptor type 5 (EC 3.1.3.48) (Protein-tyrosine phosphatase striatum-enriched) (STEP) (Neural-specific protein-tyrosine phosphatase) - Rattus norvegicus (Rat), 369 aa.	172..540 1..369	346/369 (93%) 354/369 (95%)	0.0
Q9BE09	HYPOTHETICAL 42.1 KDA PROTEIN - Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), 370 aa.	128..440 1..313	304/313 (97%) 306/313 (97%)	0.0
JC4263	protein-tyrosine-phosphatase (EC 3.1.3.48), receptor type PCPTP1 precursor - rat, 656 aa.	110..538 222..654	220/442 (49%) 286/442 (63%)	e-114

PFam analysis predicts that the NOV57a protein contains the domains shown in the Table 57E.

Table 57E. Domain Analysis of NOV57a

Pfam Domain	NOV57a Match Region	Identities/ Similarities for the Matched Region	Expect Value
Y_phosphatase	297..529	105/278 (38%) 197/278 (71%)	2.3e-103

Example 58.

The NOV58 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 58A.

Table 58A. NOV58 Sequence Analysis		
	SEQ ID NO: 169	2937 bp
NOV58a, CG97842-01 DNA Sequence	ATCAGTTCTCCACTATCCTTCTGTTTTCTAGGTAACATAACTACCCAAATATGCGAG CCATGGAGTCCATGCTTAATAAATTGAAGAGTACTGTTACAAAAGTAACAGCTGATGT CACTAGTGTCTGTAATGGGAAATCCTGTCACTAGAGAAATTTGATGTTGGTCGACACATT GCCAGTGGTGGCAATGGGCTAGCTTGGGAAGATTTTAAATGGCACAAAAAGTCAACAA AGCAGGAAGTGGCAGTTTTTGTCTTTGATAAAAACTGATTGACAAGTATCAAAAATT TGAAAAGGATCAAATCATTGATTCTCTAAAACGAGGAGTCCAACAGTTAACTCGGCTT CGACACCTCGACTTCTTACTGTCCAGCATCCTTTAGAAGAAATCCAGGGATTGCTTGG CATTTGTACAGAACCAAGTTTTTGCAGTTTAGCCAATGTTCTTGGTAACCTGGGAAAA TCTACCTTCCCTATATCTCCAGACATTAAGGATTATAAACTTTATGATGTAGAAACC AAATATGGTTTGTCTCAGGTTTCTGAAGGATTGTCTTCTGTCATAGCAGTGTGAAAA TGGTGCATGGAAATATCACTCCTGAAAATATAATTTGAATAAAAGTGGAGCCTGGAA AATAATGGGTTTTGATTTTGTGTATCATCAACCAATCCTTCTGAACAAGAGCCTAAA TTTCTTGTAAAGAATGGGACCCAAATTACCTTCATTGTGTCTTCCAAATCCTGAAT ATTTGGCTCCTGAATACATACTTCTGTGAGCTGTGAAACAGCCAGTGTATGTATTCT TTTAGGAACGTGTATGTATGCTGTATTAAATAAAGGAAACCTATATTGAAGTCAAC AAGCAAGATATTTACAAGAGTTTCAGTAGGCAGTTGGATCAGTTCAGTTCGTTTAGGAT CTAGTTCACTTACAATATACCTGAGGAAGTTCGTGAACATGTAAGCTACTGTTAAA TGTAACCTCCGACTGTAAGACCAGATGCAGATCAAATGACAAGATTCCCTTCTTTGAT GATGTTGGTGCACTAACACTGCAATATTTTGATACCTTATTTCCAAAGAGATAATCTTC AGAAATCACAGTTTTTCAAAGGACTGCCAAAGGTTCTACCAAACTGCCCAAGCGTGT CATTGTGACAGAAATTTGCCTTGTGTGACTTCAGAAATTTGTAACCCCTGACATGGTA CCTTTGTTTTGCCCAATGTTCTACTTATTGCTGAGGAATGCACCAAGAAGAATATG TCAAAATTAATTTCTCCGAACTTGGCCCTGTGTTTAAAGCAGCAGGAGCCAATCCAGAT TTTGTTAATTTTCTACAAAAATGGATTGTCTACTAACCAAAACCCCTCCTGATGAG ATAAAGAACAGTGTCTACCCATGGTTTACAGAGCACTAGAAGCTCCTTCCATTTCAGA TCCAGGAGCTCTGTCTAAACATCATTCCAACCTTTGCAAACTCTATAGACTACCCATC CATGAAAAACGCTTTGATACCAAGAATTAATAATGCTGTCTACAAACATCTTCCCTT GCGGTTCTGTGTAATTCATTAGTGTGCTTAGGAAGATTTTGGAACTACTTGGATAAGT GGTTCCTCATGGGAATTTTAGGTATTTACAAATGTACTTTTACTCATAAGAAGTTGGGA ATCACCAGAGCAGCTGGCCGAAAAGTGTGCCTCATCTTATCCCTGAGTATTG AAAACAATCTTAATCTTAATCAGTTCAATCTTTCAATTCCTGATATAAAGAAATGCT TAATAGATTGGAGTCTGAACATAAGACTAACTGGAGCAACTTCATATAATGCAAGAA CAGCAGAAATCTTTGGATATAGGAATCAAATGAATGTTTCTGAGGAGATGAAAGTTA CAAATATTGGGAATCAGCAAATGACAAAGTTTTTAACAACATTGGAGCAGACCTTCT GACTGGCAGTGAAGTCCGAAAATAAAGAGGACGGGTACAGAATAAACATAAAGAGCA TCACTTACACTTGAAGAAAAACAAAAATTAGCAAAAGAACAGAGCAGGCACAGAAGC TGAAAAGCCAGCAGCCTCTTAAACCCCAAGTGCACACACCTGTTGCTACTGTTAAACA GACTAAGGACTTGACAGACACACTGATGGATAATATGTCATCCTTGACAGCCTTTCT GTTAGTACCCCTAAATCTTCTGCTTCAAGTACTTTCACTTCTGTTCTTCCATGGGCA TTGGTATGATGTTTTCTACACCACTGATAATACAAAGAGAAATTTGACAAATGGCCT AAATGCCAATATGGGCTTTCAGACTTCAGGATTCAACATGCCCGTTAATACAAACAG AACTTCTACAGTAGTCCAAGCAGTTGGAGTGACCAAGATGACTCTGGGAACACCTC CCCTTTGCCAACTTCAATGCTTTGAGTGTTCCTCTGTGCTGGTCAAGCAGACCCA ACAAGACCCACAGATATGTCTGCCCTTAATAATCTCTTTGGCCCTCAGAAACCCAAA GTTAGCATGAACAGTTATCACAACAGAAACCAAATCAGTGGCTTAATCAGTTGTAC CTCCTCAAGGTTCTCCAATATGGGCAGTTTCAATATGGGACACAGATGAACGTGAT AGGACAATCTGCTTTTGGTATGCAGGTAATCCTTCTTAAACCCACAGAATTTGCA CAGCCACCACTACTATGACCAATAGCAGTTTCACTAGCAATGATTTAAAGATCTTT TTGGGTGAGGTGCTTACTTCTATTTGAAGGATTATTTCAGTTTCAATCATGGGTGA GCTGATTACATCTTTATATAGTTGGCTTGGAGGAAG	

	ORF Start: ATG at 61	ORF Stop: TGA at 2848	
	SEQ ID NO: 170	929 aa	MW at 103707.8kD
NOV58a, CG97842-01 Protein Sequence	MESMLNKLKSTVTTKVTADVTSAVMGNPVTREFDVGRHIASGGNGLAWKIFNGTKKSTK QEVAVFVFDKKLIDKYQKFEKDQIIDSLKRGVQQLTRLRHPRLLTVOHPLEESRDCLA FCTEPVFASLANVLGNWENLPSPISPDIKDYKLYDVETKYGLLQVSEGLSFLHSSVKM VHGNITPENIILNKSGAWKIMGFDFCVSSTNPSEQEPKFPCKEWDPNLPSLCLPNPEY LAPEYILSVSCETASDMYSLGTVMYAVFNKGKPIFEVNKQDIYKSFSRQLDQLSRLGS SSLTNIPEEVREHVKLLNVTPTVRPDADQMTKIPFFDDVGAVTLQYFDTLRFQNDNLQ KSQFFKGLPKVLPKLPKRIVVQRIPLCLTSEFVNPDMPVFLPNVLLIABECTKEEYV KLILPELGPVFKQEQPIQILLIFLQKMDLLLTTPPDEIKNSVLPVRYRALEAPSIQI QELCLNIIPTFANLIDYPSMKNALIPRIKNACLQTS SLAVRVNSLVCGLKILEYLDKW FVLDDILPFLQQIPSKPAVLGMILGIYKCTFTHKKLGITKEQLAGKVLPHLIPLSIE NNLNLNQFNSFISVIKEMLNRLSEHKT KLEQLHIMQEQQSLDIGNQMNVSSEMKVT NIGNQQIDKVFNNIGADLLTGSESENKEDGLQNKHKRASLTLEEKQKLAKQEQQAQKL KSQQPLKPQVHTPVATVQTKDLTDTLMDNMSSLTSLSVSTPKSSASSTFTSVPSMGI GMMFSTPTDNTKRLTNGLNANMGFQTSGFNMFPVNTNQNFYSSPSTVGVTKMTLGTTP TLPNFNALSVPAPAGAKQTQQRPTDMSALNNLFGPQKPKVSMNQLSQQKPNQWLNQFVP PQGSPTMGSSVMGTQMNVIQSAFGMQGNPFFNPQNFAQPPTMTNSSSASNDLKDLF G		

Further analysis of the NOV58a protein yielded the following properties shown in Table 58B.

Table 58B. Protein Sequence Properties NOV58a	
PSort analysis:	0.4564 probability located in mitochondrial matrix space; 0.3000 probability located in microbody (peroxisome); 0.3000 probability located in nucleus; 0.1507 probability located in mitochondrial inner membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV58a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 58C.

Table 58C. Geneseq Results for NOV58a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV58a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB85782	Human kinase PKIN-1 - Homo sapiens, 792 aa. [WO200160991-A2, 23-AUG-2001]	153..929 12..792	771/781 (98%) 772/781 (98%)	0.0
AAG67427	Amino acid sequence of a human protein kinase/protein phosphatase - Homo sapiens, 756 aa. [WO200109345-A1, 08-FEB-2001]	174..929 1..756	755/756 (99%) 755/756 (99%)	0.0
AAB95177	Human protein sequence SEQ ID NO:17239 - Homo sapiens, 756 aa. [EP1074617-A2, 07-FEB-2001]	174..929 1..756	755/756 (99%) 755/756 (99%)	0.0
AAB95772	Human protein sequence SEQ ID NO:18711 - Homo sapiens, 735 aa. [EP1074617-A2, 07-FEB-2001]	195..929 1..735	732/735 (99%) 734/735 (99%)	0.0
AAB65678	Novel protein kinase, SEQ ID NO: 206 - Homo sapiens, 505 aa. [WO200073469-A2, 07-DEC-2000]	1..496 1..496	494/496 (99%) 494/496 (99%)	0.0

In a BLAST search of public sequence databases, the NOV58a protein was found to have homology to the proteins shown in the BLASTP data in Table 58D.

Table 58D. Public BLASTP Results for NOV58a				
Protein Accession Number	Protein/Organism/Length	NOV58a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9P2I7	KIAA1360 PROTEIN - Homo sapiens (Human), 796 aa (fragment).	138..929 1..796	788/796 (98%) 789/796 (98%)	0.0
Q96ST4	CDNA FLJ14645 FIS, CLONE NT2RP2001839, WEAKLY SIMILAR TO SCY1 PROTEIN - Homo sapiens (Human), 756 aa.	174..929 1..756	755/756 (99%) 755/756 (99%)	0.0
Q9H7V5	CDNA FLJ14212 FIS, CLONE NT2RP3003500, WEAKLY SIMILAR TO SCY1 PROTEIN - Homo sapiens (Human), 735 aa.	195..929 1..735	732/735 (99%) 734/735 (99%)	0.0
Q9NVH3	CDNA FLJ10735 FIS, CLONE NT2RP3001407, WEAKLY SIMILAR TO SCY1 PROTEIN - Homo sapiens (Human), 446 aa.	484..929 1..446	445/446 (99%) 445/446 (99%)	0.0
Q96EF4	UNKNOWN (PROTEIN FOR MGC:8814) - Homo sapiens (Human), 446 aa.	484..929 1..446	444/446 (99%) 444/446 (99%)	0.0

Pfam analysis predicts that the NOV58a protein contains the domains shown in the Table 58E.

Table 58E. Domain Analysis of NOV58a			
Pfam Domain	NOV58a Match Region	Identities/ Similarities for the Matched Region	Expect Value
pkinese	159..327	44/203 (22%) 116/203 (57%)	3e-07

Example 59.

- 5 The NOV59 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 59A.

Table 59A. NOV59 Sequence Analysis

	SEQ ID NO: 171	2933 bp	
NOV59a, CG98021-01 DNA Sequence	ACTCTATGTCCTCTCGTTGGATTGTGACACCGGGAGGTGAGGAACTCCAGGACCT TGTTCCTGCTGGATTGCGAGCAACAGCACAGCACGTAGGGCGTAGTTGGTGCTGG TGGATGTTTGTGAATGAATGAATGATGAATGGCTGGCACCTTGTCTGCTCATCCCTA ACTCCTGTTCCCTTCATCTGTGCAGCCCTAATCTTTGTTTCTCATCTGTCCATCCCT TATTTGTGCATCCTCATTCTTAGCCCCCTTCACTGCCCTTCTCCATCTCTTCTCCTTG TTCATTTGTCCCTGTTCTCTGTCTCTACTCCACTCATGCCATCTCTGTGCCCTTGA CTTACCCAGTCCCTGCTACTATCTCCATCCCTAATTTCTGCCCTCTTGTCTGTCTACT CCTAATTCTTTTCTTGTCCATCCCTAATACCTGTACCTTGTCTCTTCTGGAAGTTGTG TCTCCATCCCTAATCCATCTGCCCTAATCTCTGTGCCCTTTGCCCATCCTTCTCTTT CTCGGTTGTCTCTTTCCACCCCTTATCTCCACACCTGCCACCCCTGCACTCCCATCTGT TTCCCATCTGCACCCCTTGCCCATCCCTCCACACACAGGACAGACGGCCACCATGT CAGGAGACTACGAGGATGACCTCTGCCGGCGGGCACTCATCTGGTCTCGGACCTCTG TGCGCGGGTCCGAGATGCTGACACCAACGACAGGTGCCAGGAGTTCAATGACCGAATC CGAGGCTATCCCGGGGTCCAGATGCAGACATCTCCGTGAGCCTGCTGTCTGGTCTATCG TGACATTCTGTGGCATTGTCTTCTGGGTGTCTCTCTCTCTGTGTCTGTGAAGTTGTG CTGGGTGCCCTGGCGGGACAAGGAGGCTCGGCAGTGGGCGGTGGCCCCCTGCGCAA GACCTAGGCCCTGTGTCTGGGCTGGCAGGCCTGGTAGGCGGAGCGGGACCATCTGG CGGCTGGCCTGGGTGGCCATCCTCTGCTGGGCGGCCACACCATATGCCATGGCGC CCACCATCCACCTTTGTGTAGCTGTGTGGAGCCAGGCAGCTGGGGGGTTCTGACACC CCTGAGCCCTCTACTTGGACATGGACTCGTATCCAGAGGCTGCAGCAGCAGCAGTGG CCGCTGGGGTCAAACCGAGCCAAACATCCCTGAGCTGCCCTCTGAGGGGGGAGCAGG CTCTGGGTGTCTCTCTGTGCCCCCAAGTGGTGGGGGCTTGGCCAGTGCAGTGCATCA CAGCAGGTCAAGCCTGGCACCCTACCAGGTACCAGCCCTGCCCGACCCCTCA CCAGCAGACTCTGACCTCCAGCCGGAACCCAGCAGTGAGGAGCGCCACCTGCCCT GCCCTTACCCCTGCTTGGAGGCGAGGAAAAAGCCAACTCATTGGGAGATTAAAGCCA GAGCTGTACAGGGGAGCTGGCCCTGGTGGCCGGCGGAGCGGTGGGGGCCAGGCTCTG GAGAGGCAGGCACAGGGGCACCTGTGGCCGTATCAGCTTGGCCCTGCGGTACCTCTA TGGCTCGGACCAGCTGGTGGTGGAGATCTGCAGGCCCTGGACCTCCCTGCCAAGGAC TCCAAAGGCTTCTCAGACCCCTACGTCAAGATCTACCTGTGCTGACCGCAAGAAAA AGTTTCAGACCAAGGTGCACAGGAAGACCCCTGAACCCCGTCTTCAATGAGACGTTTCA ATTCTCGGTGCCCTTGGCCGAGCTGGCCCAACGCAAACTGCATCTCAGCGTCTATGAC TTTGACCGCTTCTCGCGGCACGACCTCATCGGCCAGGTGGTGTGGACAACCTCCTGTG AGCTGGCCGAGCAGCCCCCTGACCGCCCGCTCTGGAGGGACATCGTGGAGGGCGCTC GGAAAAAGCAGATCTTGGGGAGCTCACTTCTCACTCTGCTACCTCCCAAGCGCGG CGCCTCACCCTGACCATCATCAAAGCCTCTAACTCAAAGCGATGGACCTCACTGGCT TCTCAGACCCCTACGTGAAGGCCTCCCTGATCAGCGAGGGGCGGCGTCTGAAGAAGCG GAAAACCTCCATCAAGAAGAACGCTGAACCCACCTATAATGAGGCGTGGTGTTC GACGTGGCCCCCGAGAGCGTGGAGAACGTGGGGCTCAGCATCGCCGTGGTGGACTACG ACTGCATCGGGCACAACGAGGTGATCGGCGTGTGCCGTGTGGGCCCGGACGCTGCCG CCCGACGGCCCGGAGCACTGGGCAGAGATGCTGGCCAATCCCCGCAAGCCCGTGGAG CACTGGCATCAGCTAGTGGAGGAAAAGACTGTGACCACTTCAAAAAGGCAGCAAAG GACTATCAGAGAAAGAGAACTCCGAGTGAAGGGTCTGGCCTAGGCCCCGGGATCCGAC AGGCTCCCTCAGGACCCCATCTTCTCTGCCCGGACCGTGAATTCATCTCTTGAAGC CATAACGTCGAGCTGCTGGTGGGGGCGAGCCCTGGCCCTAGGCTTCTCAACCTTGG AGCGAGAGGATGAGAGGAGGCGGCCAGCTCCTTCTTTCAGGGTGGGGGTCACTCAG CCTCCACTGTGTCTGTCTTTCTTCTCCCTGGGGCTGCCCTCGAGGCGAGGGGCCATG ATGTCTGGGGGACCCCTGCCCCCAAACCCCTCTGTCTGTCTCTGTCTTGTGTGT TGTCCAAGACTCAGTGTCCCGACCCCTGTCTCGCCGTGAATGTCAATGGGCCAATC TCTCTGTCTTTTTCAGACACACACACCTGTGTCCACCCCTTCTGTTCGCCACACCC GCGTCTGGCCGGTCCCCCACTGCTGTCTGTATCAACGCCAGAATAAACACACTCTGT GGGTCTCACTCCAAAAAAGAAAAAAGC		
	ORF Start: ATG at 635	ORF Stop: TGA at 2405	
	SEQ ID NO: 172	590 aa	MW at 63303.0kD
NOV59a, CG98021-01 Protein Sequence	MSGDYEDDLRRLILVSDLCARVRDADTNDRCQEFNDRIRGYPRGPDADISVLLSV IVTFPGVILLGVSLFVSWKLCWPWRDKGSAVGGGPLRKDLGPGVGLAGLVGGGGHH LAAGLGGHPLLGPHHHAAHHPFAELLEPGSLGSDTPEPSYLDMDSYPEAAAAA VAAGVKPSQTSPELPSEGGAGSGLLLLPPSGGGLPSAQSHQVTSAPTRYPALPRP LTQQLTSLQPDPSSEERPPALPLPLPGGEKAKLIGQIKPELYQGTGPGGRRSGGGP SGEAGTGAPCGRISFALRYLYGSDQLVVRILQALDLPKDSNGFSDPYVKIYLLPDRK KKFQTKVHRKTLNPFVNETFQFSVPLAELAQRLHFSVYDFDRFSRHLIGQVVLNLL LELAEQPPDRPLWRDIVEGGSEKADLGLNLSLYLPTAGRLTVTIKASNLKAMDLT GFSDDPYVKASLISEGRLKRRKTSIKKNTLNPTYNEALVPDAPESVENVGLSTAVVD YDCIGHNEVIGVCRVGPDAADPHGREHWAEMLANPRKPVEHWHQLVEEKTVTSPTKGS KGLSEKENSE		

Further analysis of the NOV59a protein yielded the following properties shown in Table 59B.

Table 59B. Protein Sequence Properties NOV59a	
PSort analysis:	0.6000 probability located in endoplasmic reticulum (membrane); 0.4600 probability located in plasma membrane; 0.3128 probability located in microbody (peroxisome); 0.3000 probability located in nucleus
SignalP analysis:	Cleavage site between residues 28 and 29

A search of the NOV59a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
5 several homologous proteins shown in Table 59C.

Table 59C. Geneseq Results for NOV59a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV59a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU19714	Human novel extracellular matrix protein, Seq ID No 364 - Homo sapiens, 295 aa. [WO200155368-A1, 02-AUG-2001]	291..571 1..281	278/281 (98%) 278/281 (98%)	e-161
AAU19715	Human novel extracellular matrix protein, Seq ID No 365 - Homo sapiens, 461 aa. [WO200155368-A1, 02-AUG-2001]	77..583 33..461	235/509 (46%) 304/509 (59%)	e-116
AAU19847	Human novel extracellular matrix protein, Seq ID No 497 - Homo sapiens, 208 aa. [WO200155368-A1, 02-AUG-2001]	309..479 19..189	161/171 (94%) 162/171 (94%)	9e-88
AAR97722	Mouse inositol polyphosphate binding protein IP4-BP - Mus musculus, 422 aa. [JP08092290-A, 09-APR-1996]	288..581 130..422	148/297 (49%) 203/297 (67%)	4e-76
ABB59660	Drosophila melanogaster polypeptide SEQ ID NO 5772 - Drosophila melanogaster, 474 aa. [WO200171042-A2, 27-SEP-2001]	301..581 194..474	148/287 (51%) 200/287 (69%)	1e-75

In a BLAST search of public sequence databases, the NOV59a protein was found to have homology to the proteins shown in the BLASTP data in Table 59D.

Table 59D. Public BLASTP Results for NOV59a				
Protein Accession Number	Protein/Organism/Length	NOV59a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9BQG1	Synaptotagmin III (SytIII) - Homo sapiens (Human), 590 aa.	1..590 1..590	590/590 (100%) 590/590 (100%)	0.0
Q925B7	SYNAPTOTAGMIN 3 - Rattus norvegicus (Rat), 588 aa.	1..590 1..588	566/592 (95%) 576/592 (96%)	0.0
P40748	Synaptotagmin III (SytIII) - Rattus norvegicus (Rat), 588 aa.	1..590 1..588	564/592 (95%) 575/592 (96%)	0.0
O35681	Synaptotagmin III (SytIII) - Mus musculus (Mouse), 587 aa.	1..590 1..587	561/592 (94%) 572/592 (95%)	0.0
P24507	Synaptotagmin C (Synaptic vesicle protein O-P65-C) - Discopyge ommata (Electric ray), 537 aa.	1..576 1..512	342/578 (59%) 405/578 (69%)	0.0

Pfam analysis predicts that the NOV59a protein contains the domains shown in the Table 59E.

Table 59E. Domain Analysis of NOV59a			
Pfam Domain	NOV59a Match Region	Identities/ Similarities for the Matched Region	Expect Value
C2	316..402	56/97 (58%) 82/97 (85%)	1.4e-44
C2	448..536	41/97 (42%) 79/97 (81%)	1.5e-38

Example 60.

- 5 The NOV60 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 60A.

Table 60A. NOV60 Sequence Analysis

Table 60A. NOV60 Sequence Analysis			
	SEQ ID NO: 173		2133 bp
NOV60a, CG98030-01 DNA Sequence	CGTTTACGGCATGTCATTCTGGACACATGGCATGTTCCATGGCGTGTGGCGGTAGAG CTGCAAGTATGAGAACCCAGCCGCTGGAGTGAAGCAGGAGCAAGCCATTAAAGGGGGT TTAATCATCTGGGTCACTGATAATATACTGGCCATGGCCCGCCCATCTCTGAGCTC CTGGAGAAGTACCACATCATTGATCAGTTCCTCAGCCATGGCATAAAAACAATAATCA ACCTCCAGCGCCCTGGTGAGCATGCTAGCTGTGGGAACCCCTCTGGAACAAGAAAGTGG CTTACATACCTTCTGAGGCTTTATGAGGCTGGCATTACTTCTACAATTTGCGGA TGGAAAGGATTATGTTGATGCTCTTACTACTATCTTAGATATGGTGAAGGTGATGA CATTTCGCTTACAGGAAGGAAAGTAGCTATCCATTGTGATGCAGGGCTTGGTCGAAC AGGTGTTTTAATAGCTTGTACTTAGTTTTTGCAACGAGAATGACTGCTGACCAAGCA ATTATATTTGTGCGGGCAAAGCGACCCAAATCCATACAAACCAAGAGACAGCTCTCT GTGTAAGGGAATTTACTCAGTTTCTAACTCCTCTCCGCAATATATCTCTTGCTGTGA TCCCAAGCACATGCTGTACCTTACCTCAATATCTAATTCGCCAGCGTCATCTGCTT CATGTTATGAGGACGACTTCTGAAACACGTGCCAAAAATTATCCACCTAGTTTGCA AATTGCTGCTGGACTTAGCGGAGAACAGGCCAGTGATGATGAAGGATGTGTCCGAAGG ACCTGGTCTCTCTGCTGAAATAGAAAAGACAAATGTCTGAGATGGTCACCATGCAGCTG GATAAAGAGTTACTGAGGCATGACAGTGATGTGTCCAACCCGCTAACCCACTGCGAG TGGCAGCAGATTTTGACAATCGAGGCATGATTTTCTCCAATGAGCAACAGTTTGACCC TCTTTGGAAGAGCGGAATGTTGAGTGCCCTTCAACCCCTGACTCATCTGAAAGGCGG CTCAGCTACAGTGACTCAGATTTAAAGAGGGCCGAGAACCCTCTGGAGCAAGGGGAGA CTCCACAGACAGTGCCTGCCAGATCTTGGTTGGCCACAAGCCAGCCAGCAGAAGCT CATAAGCCATTGTTACATCCCAGTCTCCAGAACCAGACTTACACAAGGAAGCCTTG GTTCGACGACACTTTCTTCTGGAGTCAGTCAAAGTTTGGAGGCTGGAAGGACTCA AAGATAATGGGTCAACATTTTCCATGGAAGGATCATTCCAAAGGAAGCACAGCAGAG TGGAGCTTTCTCTGCAGATGTTTTCAGGCTCACACAGCCCTGGGGAGCCAGTTTCAACC AGCTTTGCAAAATGCCATAAGGATCCAAACCTGCTCACCAGCAAGTGTCTCACTGTC AGTGTAAGAACTCATGGTGTGGGAGCCCTGGCTCTGTGAGGCAGAACAGCAGGACACC CCGAAGCCCTCTGGACTGTGGCTCCAGTCCAAAGCACAGTTCTTGGTTGAACATGAA ACCCAGGACAGTAAAGATCTGTCTGAAGCAGCTTCACTCTGCATTACAGTCTGAAT TGAGTGCTGAGGCAAGAAGAACTAGTGGCGGCCAAAGCCCTAGCAAAATTAAATGAATC TGTAAGAAAGGAGGAACTAAAAAGGAAGAAAGAGCTTAATCCCGAGATGGAGCTTGG GAAAGAATATGTGGCGAGAGGAGCCCTTTCATCCTATGCAGCTTGTGTTGCTTGGG TGGAGCAACTGAAGGAGCCTGTAATCACCAGAGGATGTGGACATGTTGGTTGACAG GCGAGCAGATGCCGAGAGCACTTTTATTAGAGAAGGGAGCAGCAGACTATT CTCTGCGTGTGTCAGTGCATAGTGAACCTGCAGACAATCCCGTGGATGTGGAGGAAG CTTTCCTTGCCCATGCCATTAAAGGCATTCACTAAGGTTAATTTTGATTCTGAAATGG ACCAACAGTTTACAACACCCTGAAGAAAATATTTAAGCACAGCTGGAAGAAAAAGA AAAATGACAAAAGATGGCCCTAAGCCTGGCCTCTAGCTTCACTC		
	ORF Start: ATG at 28		ORF Stop: TAG at 2122
	SEQ ID NO: 174		698 aa MW at 78224.7kD
NOV60a, CG98030-01 Protein Sequence	MACSMACGGGRACKYENPARWSEQEQAIGVYSSWVTDNILAMARPSSELLEYHIIDQ FLSHGIIKTIIINLQRPGEHASCGNPLEQESGFTYLPFAFMEAGIYFYNGWKDYGVASL TTILDMVKVMTFALQEGKVAIHCHAGLGRGTGLIACYLVFATRMADQAIIFVRAKRP NSIQTKRQLLCVREFTQFLTPLRNIFSCDDPKAHAVTLPOYLIRQRHLLHGYEARLLK HVPKIIHLVCKLLLDLAENRPVMMKDVSEGPGLSAEIEKTMSEVMVTLQDKELLRHDS DVSNPPNPTAADAADFNRMGIFSNQQFDPLWKRNVCEQLPLTHLKRRLSYSDSLK RAENLLEQGETPQTVPAQILVGHKPRQQLISHCYIPQSPPEPLHLKEALVRSLSFWS QSKFGGLEGLKONGSPIFHGRIIPKEAQSGAFSADVSGSHSPGEPVSPSFANVHKDP NPAHQQVSHQCCKTHGVGSPGSRVQNSRTPRSPLDCGSSPKAQFLVEHETQDSKDLSE AASHSALQSELSAEARRILAAKALANLNESEVEKEELKRKELNSRDGAWERICGERDP FILCSLMWSWVEQLKEPVIKEDVDMLVDRRAAEALFLEKGGHQTILCVLHCIVN LQTIPLVDVEEFLAHAIKFTKVNFDSENGPTVYNTLKKIFKHTLEKRKMTKDGPKP GL		

	SEQ ID NO: 175	2147 bp	
NOV60b, CG98030-02 DNA Sequence	CGTTTACGGCATGTTCATTCTGGACACATGGCATGTTCCATGGCGTGTGGCGGTAGAG CTTGC AAGTATGAGAACCAGCCCGCTGGAGTGAGCAGGAGCAGCCATTAAAGGGGT TTACTCATCTCTGGGTCACTGATAATATACTGGCCATGGCCCGCCCATCTCTGAGCTC CTGGAGAAGTACCACATCATTGATCAGTTCCTCAGCCATGGCATAAAAAACAATAATCA ACCTCCAGCGCCCTGGTGAGCATGCTAGCTGTGGGAACCTCTGGAACAAGAAAGTGG CTTCACATACCTTCTCTGAGGCTTTCATGGAGGCTGGCATTACTTCTACAATTTCGGA TGGAAGGATTATGGGTAGCGTCTCTTACTACTATCCTAGATATGGTGAAGGTGATGA CATTGCGCTTACAGGAAGGAAAAGTAGCTATCCATTGTGCATGCAGGGCTTGGTTCGAAC AGGTGTTTTAATAGCCTGTTACTTTAGTTTTTGC AACGAGAATGACTGCTGACCAAGCA ATTATATTGTGCGGGCAAAGCGACCCAATTCCATACAAACAGAGGACAGCTCCTCT GTGTAAGGGAATTTACTCAGTTTCTAACTCCTCTCGCAATATATTCTCTGCTGTGA TCCCAAAGCACATGCTGTCACTTACCTCAATATCTAATTCCGACGCGTCTATCTGCTT CATGGTTATGAGGCACGACTTCTGAAACACGTGCCAAAAAATTATCCACCTAGTTTGA AATTGCTGCTGGACTTAGCGGAGAACAGGCCAGTGATGATGAAGGATGTGTCCGAAGG ACCTGGTCTCTCTGCTGAAATAGAAAAGACAATGTCTGAGATGGTCACCATGACGCTG GATAAAGAGTTACTGAGGCATGACAGTGATGTGTCCAAACCCGCTAACCCCACTGCAG TGGCAGCAGATTTTGACAATCGAGGCATGATTTTCTCCAATGAGCAACAGTTTGAAGG TCTTTGGAAAAGGCGGAATGTTGAGTGCCCTCAACCCCTGACTCATCTGAAAAGGCGG CTCAGCTACAGTGACTCAGATTAAAGAGGGCCGAGAACCTCTGGAGCAAGGGGAGA CTCCACAGACAGTGCCCTGCCAGATCTTGGTTGGCCACAAGCCAGGCAGCAGAAGCT CATAAGCCATTGTTACATCCACAGTCTCCAGAACAGACTTACACAAGGAAGCCTTG GTTGCGAGCACACTTTCTTCTGGAGTCAGTCAAAGTTTGGAGGCTTGGAGGACTCA AAGATAATGGGTACCAATTTCCATGGAAGGATCATTCCAAGGAAGCACAGCAGAG TGGAGCTTTCTCTGCAGATGTTTCAGGCTCACACAGCCCTGGGAGCCAGTTTCAACC AGCTTTGCAAAATGTCCATAAGGATCCAAACCCCTGCTCACCAGCAAGTGTCTCACTGTC AGTGTAAGAACTCATGGTGTTGGGAGCCCTGGCTCTGTGAGGAGAACAGCAGGACACC CCGAAGCCCTCTGGACTGTGGCTCCAGTCCCAAAGCACAGTTCTTGGTTGAACATGAA ACCCAGGACAGTAAAGATCTGTCTGAAGCAGCTTCACTCTGCATTACAGTCTGAAT TGAGTGCTGAGGCAAGAAGAATACTGGCGGCCAAAGCCCTAGCAATTTAAATGAATC TGTAAGAAAAGGAGGAATAAAAAAGGAAGGTAGAAATGTGGCAGAAAGAGCTTAATTCC CGAGATGGAGCTTGGGAAGAATATGTGGCGAGAGGGACCCCTTTCATCTATGCAGCT TGATGTGGTCTTGGGTGGAGCAACTGAAGGAGCCTGTAATCACCAGAGGATGTGGA CATGTTGGTTGACAGGCGAGCAGATGCCGAGAAGCACTTTTTTTATTAGAGAAGGGA CAGCACCAGACTATTCTCTGCGTGTGTCACTGCATAGTGAACCTGCAGACAATTCCCG TGGATGTGGAGGAAGCTTTCTTGCCCATGCCATTAAAGGCATTCAAGGTTAGTTT TGATTCTGAAAATGGACCAACAGTTTACAACACCCTGAAGAAAAATTTAAGCACACG CTGGAAGAAAAAGAAAATGACAAAAGATGGCCCTAAGCCTGGCCTCTAGCTTTTAC T		
	ORF Start: ATG at 28	ORF Stop: TAG at 2137	
	SEQ ID NO: 176	703 aa	MW at 78800.4kD
NOV60b, CG98030-02 Protein Sequence	MACSMACGGGRACKYENPARWSEQEQAIGVYSSWVTDNILAMARPSSELLEKYHIIDQ FLSHGIKTIINLQRPGEHASCGNPLEQESGFTYLPFAFMEAGIYFYNFGWKDYGVASL TTILDMVKVMTFALQEGKVAIHCHAGLGRGTGLIACYLVFATRMTADQAIIFVRAKRP NSIQTRGQLLCVREFTQFLTPLRNI FSCCDPKAHAVTLPQYLIRQRHLLHGYEARLLK HVPKIIHLVCKLLDLAENRPVMMKDVSEGPGLSAEIEKTMSEVMVTMQLDKELLRHDS DVSNNPPNPTAVAADFDRGMIFSNQEQFDPLWKRRNVECLQPLTHLKRRLSYSDSDLK RAENLLEQGETPQTVPAQILVGHKPRQQLISHCYIPQSPEPDHKEALVRSSTLSFWS QSKFGGLEGLKDNNGSPIFHGRIIPKEAQSGAFSADVSGSHSPGEPVSPSFANVHKDP NPAHQVQVSHCQCKTHGVGSPGSRVQNSRTPRSPDCGSSPKAQFLVEHETQDSKDLSE AASHSALQSELSAEARRILAAKALANLINESVEKEELKRKVMWQKELNSRDGAWERIC GERDPFIIICSLMWSWVEQLKEPVI TKEDVDMLVDRRADAAEALFLEKGGHQHTILCVL HCIVNLQTI PVDVEEAF LAHAIFAFTKVSFDSSENGPTVYNTLKKI FKHTLEKRMKTK DGPKPGL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 60B.

Table 60B. Comparison of NOV60a against NOV60b.		
Protein Sequence	NOV60a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV60b	1..698 1..703	695/703 (98%) 697/703 (98%)

Further analysis of the NOV60a protein yielded the following properties shown in Table 60C.

Table 60C. Protein Sequence Properties NOV60a	
PSort analysis:	0.9400 probability located in nucleus; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV60a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 60D.

Table 60D. Geneseq Results for NOV60a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV60a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG08740	Novel human diagnostic protein #8731 - Homo sapiens, 476 aa. [WO200175067-A2, 11-OCT-2001]	432..662 227..462	229/236 (97%) 231/236 (97%)	e-129
ABG08740	Novel human diagnostic protein #8731 - Homo sapiens, 476 aa. [WO200175067-A2, 11-OCT-2001]	432..662 227..462	229/236 (97%) 231/236 (97%)	e-129
AAG67632	Amino acid sequence of a human protein - Homo sapiens, 447 aa. [WO200109316-A1, 08-FEB-2001]	16..190 171..335	53/183 (28%) 79/183 (42%)	6e-10
AAG67453	Amino acid sequence of a human polypeptide - Homo sapiens, 447 aa. [WO200109345-A1, 08-FEB-2001]	16..190 171..335	53/183 (28%) 79/183 (42%)	6e-10
AAU75364	Human dual specificity phosphatase CDC14A L572P variant - Homo sapiens, 594 aa. [US6331614-B1, 18-DEC-2001]	91..177 232..316	32/87 (36%) 46/87 (52%)	3e-07

In a BLAST search of public sequence databases, the NOV60a protein was found to have homology to the proteins shown in the BLASTP data in Table 60E.

Table 60E. Public BLASTP Results for NOV60a				
Protein Accession Number	Protein/Organism/Length	NOV60a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q8WX19	BA490F3.2.1 (NOVEL PROTEIN ISOFORM 1) - Homo sapiens (Human), 698 aa.	1..698 1..698	696/698 (99%) 697/698 (99%)	0.0
Q9N091	UNNAMED PROTEIN PRODUCT - Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), 683 aa.	1..682 1..682	654/682 (95%) 663/682 (96%)	0.0
Q8WX18	BA490F3.2.2 (NOVEL PROTEIN ISOFORM 2) - Homo sapiens (Human), 112 aa.	587..698 1..112	112/112 (100%) 112/112 (100%)	1e-59
Q9UAX0	T12B3.1 PROTEIN - Caenorhabditis elegans, 446 aa.	3..246 42..281	92/244 (37%) 147/244 (59%)	2e-44
O60730	CDC14B3 PHOSPHATASE - Homo sapiens (Human), 471 aa.	16..190 201..365	54/183 (29%) 81/183 (43%)	1e-10

Pfam analysis predicts that the NOV60a protein contains the domains shown in the Table 60F.

Table 60F. Domain Analysis of NOV60a			
Pfam Domain	NOV60a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DSPc	38..192	34/181 (19%) 98/181 (54%)	0.0011
Y_phosphatase	25..192	34/290 (12%) 105/290 (36%)	0.0046

5 Example 61.

The NOV61 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 61A.

Table 61A. NOV61 Sequence Analysis

	SEQ ID NO: 177	4172 bp
NOV61a, CG98061-01 DNA Sequence	ATGGCGGCGGCGGAGCCCTGGAACGGAGCTTCGTGGAGCTAAGCGGAGCTGAGCGCG AAAGGCCGAGGCCTTTTCGGGAATTCACAGTCTGCAGCATGGGACTGCAAAATGCCGT GGCTGGCGCCGTAAATACAGTGAAGCGCGGAGGCTTTTACTACGTGGAGAGTGGC AAGTTGTTCTCCGTAACCAAGACAGGTTTCATTTCATTGGAAGACCTCTGGAGATACAT TGGAGCTGATGGAGGAGTCACTGGACATAAATCTGTTGAATAATGCCATTTCGCCCTAAA ATTCAAAAATTGCAGTGTTTTACCTGGAGGGGTTTATGTCTCTGAGACTCAGAATCGT GTGATAATCTTGATGTTAACCAATCAAACAGTGCACAGGTTACTTTTACCACACCCCT CCCGGATGTATAGGAGTGTGTTAGTGTGACAGTGCAGTGCAGTCAATATTCAGTGA CATTGGAAGGTTGATTTCACAGATCCTTGCAACTATCAGTTAATTCAGCAGTACCT GGAATATCTCCTAATTCACCCGCTCTACAGCCTGGCTCAGCAGTGTATGGGGAGGCC TGTTCGCTTACCATGTGCTTCTGGGGGAATCTTTGTTCTTAAGTACCTCCTTATGA CATACCTGGTATGGTGTGAGTGTGGAAGTGAACAGAGTTTCAGTAATGCAACGAGTTG CTTACAGGCTGGATGCCAACAGCTATCAGGGGTGACCACTGCGCTTCAGATCGTCCCT TCAGTCTTGCTGTTTCATTGTGTGGAGCATGATGCCTTCATCTTTGCTTTGTGTGAGGA TCATAAACTACGAATGTGGTCTTACAGGAGCAAAATGTGCTTAATGGTAGCTGACATG CTGGAGTATGTCCCTGTGAAGAAAGACCTTCGGCTTACTGTGGAAGTGGACACAAAT TACGGCTTGCTTATTCCTCCACCATGGGACTCTACCTGGGGATATACATGCATGCACC AAAACGAGGACAGTTCTGCATTTTCCAGTTGGTGAGCACTGAGAGTAATCGCTATAGT CTCGATCATATTTCTTCACTGTTCACTTCTCAGGAGACACTGATTGACTTTGCTTAA CTTCCACGATATCTGGGCCCTGTGGCATGATGCTGAGAACCAACAGTAGTGAATA CATCAACTTTGAACATAATGTTGCAGGTGAGTGAATCCAGTTTTCATGCAGCCTCTG CCAGAGGAAGAGATTGTGCATCAGAGATGATCAAGACCCAGAGAGATGTATCTGCAAA GTCTTTTACACAGGACAAATTCACAAATGAAGCTTTATGTAAGGCTTTACAGATTTT CTGCCGAGGAAGTGAAGGAATTTGGATCTTTCTGGAGTGAACCTGAAGAAAGAGTT ACTTTAGCTGTGAAAATGAGCTTCAAGGAAGTGAACAGAGTATGAATCTCCCAGG AGGAGTTTCGAAATTTACAACAAGAATTCGGTGCAAGTTCTATGCTGCTTGTCTTCA GTATCAAGAAGCCCTCTCTACCCCTCTGCCCCTACATTTGAATCCACACACAAACATG GTGTGCTGCTGAAAAAAGGTTACCTGTCTTCTTATTCCTCATCTTATGTTGATC ATTTGTATCTCTGCTTATGAGAACCTTTGACAGAAGATGAGACAACCATATCTGA TGATGTGGATATCGCTCGGATGTCTATGTCTTATAAATGCCTCCGGCTGATTGAA GAGTCAGTAACCTGTGATATGTGAGTTAATGGAATGAGTTGTTATAACCTACAGT CTCCGGAAGGCTGACAGCAGATTCTGGAAGATATGATCACTATTGATGTAGAAAA TGTGATGGAGGATATTTGTAGTAACTGCAAGAGATTAGGAACCAATCCATGCAATT GGACTACTTTACGGGAATGGATTATGAAACAGAGTGAAGTGAAGGAGGATTCA ATCCAGCTCAGCCTTTGAATATTCGAATGAATCTTACCCAGCTCTATGGTAGTAACAC AGCAGGATATTTGTGTGACAGGGGTGCATAAATCGCCAGTACTCGTTTCTGTATC TGCAGAGATCTTTGATCTTACAGCAGCTGTTAATGAGGCTTGGAGATGCTGTGATTT GGGAACTGGTCAGCTCTTTCAAGCTCAGCAAGACCTACTACATCGAACCTCCCT ACTCTTATCTTATTACCTCATTAAATGGGAAGTGAAGTCTTGGCACTGATGTTCCA CTTGACACACTGGAGTCTAATCTCCAACACTTATCAGTACTGGAATTAACAGACTCTG GTGCTTTAATGGCAATAGGTTTGTATCTAGTCTCAGACTATTGTGGAGTTATCTT CCAAGAAGTTGCAAGAAAAACACATTATATCTCACCTCTTCTCAGCCAAAGGCACCT CTGAGCCAACTGGATTGAATTGGCCTGAAATGATTACTGCAATTACAGTTATTTAT TGCAGCTTTTATGGCTAGCAATCTGTTGTCTTCTTCTAGAAATGTTGATGGGAAA TTGCCAATATGTACAATTGCAGGATTATATCAACTGCTACATCCCTGGTCAAGTC AATGTTGGTCTCTGTGATTTATGCTGGGAAGGTGTTACCTAGTTACAGGAGAAGGAC AGAAGGCTCTGGAATGTTTGTGTCAGGCAGCATCTGAAGTAGGCAAGAGGAATCTT GGATCGCTTGATTCGCTCAGAGGATGGGGAGATCGTGTCTACCCCAAGGCTGCAGTAT TATGACAAGGTTTACGACTACTAGATGTCTATGGTTGCTGAACTGGTTATTTCAGT TGGCTACATCAGCCATAACTGAAGCAAGTGTGACTGGAAGTCAAGGCTACTCTAAG GACATGTATTTTCAACATCATTGGATTGGGTCACAATAGCCAGCATATGAAGCC TTAACCCAAATCTGATTCAGCAGGCAATTAGATTGTTTACGGCAGTTGGTGGTAG TTCTTTGTGAACGCTCAGACTACAGGATCTGTAGAGTTTCCCTATGTGAATCTGCA TAATGAGGTTGTGGGAATAATTGAGTCAAGTGTAGAGCTGTGGACCTTATGACTCAC AATTACTATGAACCTCTGTATGCCTTTACATCTATCGCCCAATTACCGCAAGGCTG GCACAGTGTGTTGAGTATGGAATGCGGCTTGGCAGAGAGTTCGAACCTCTCCGGG ACTTGAGAAACAAGGCAACTGTTATCTGGCTGCTCTCAATTGTTTACGACTTATTCGT CCAGAATATGCGTGAATTGTGACCCAGTGTCTGGTGCAGTGTATGATCGCCCTGGAG CATCCCTTAAGAGGAATCATGATGGAGAATGCACAGCTGCCCAAAATCGACAAAT TGAATCTCGGAACTGGGAAGATCTGGAGAAAGAGTGTCTTGGCTCGCATGCCCTC ACTTGGCTCAGCATGATCCATCAGCGGTTGCAGTTGCTGGAAGTTCATCAGCAGAGG AAATGGTCACTCTCTTGGTTTACGGCGGCTCTTTGACACTGCCATATCACTCTGTCA GACTTTTAAAGCTTCCCTTAAACGCCAGTCTTTGAAGGCTTGCCTTCAATGATCAAA TTGCAATTGGAGGAGAGGAGCAGCACAAGCAGAAGCTGGGCTGGCTAGCAGCCAATC AGCTCTCATCTGTCTACTACTAAGGAGTCTAGTGTACAGATGAAGCATGGCGACT ATTATCCACTTACCTGGAGAGGTACAAAGTCCAGAATACTTGTATCACCAGTGTGA ATCAACAAGCTCTGTCTCATGGAGTGCCTCTGCCTAATGGCTTATAAAGCAGTTACA	

	AGAAGGTTGATGCTGCTGAATGCTTCGTTTATACTTAACTATGACCTTTTAGAAGA AGCTGTGGATTGGTGTGAGAATATGTGGATGCTGATTTGGGAAAAGGACATCAATAC TTCGGAATTGAGGTAGGCATATACAAATGGTCCCAATTGGGAGAGACAGTGCCAACA GTCAACATCGCACTGTCCAGAAAATACTTGACAAATTGGAGGACTACCAGCAAAA AGTTGATAAGGCAACACGGGATTTATTATATCGTCGGACCTTGTGATTGGATT
	ORF Start: ATG at 1 ORF Stop: TGA at 4162
	SEQ ID NO: 178 1387 aa MW at 157181.6kD
NOV61a, CG98061-01 Protein Sequence	MAAAGALERSFVELSGAERERPRHFREFTVCSIGTANAVAGAVKYESAGGFYVESG KLFVTRNRFIHWKTSQDTLELMEESLDINLLNNAIRLKFQNCVLPGGVYVSETQNR VIILMLTNOTVHRLLEPHPSRMYRSELVVDSQMSIFTDIGKVDFTDPCNYQLIPAVP GISPNSTASTAWLSSDGEALFALPCASGGIFVLKLPYDIPGMVSVVELKQSSVMQRL LTGWMPTAIRGDQSPDRPLSLAVHCEHDAFI FALCQDHLKLRMWSYKEQMCLMVA LEYVPVKDLRLTAGTGHLRLAYSPTMGLYLGIMHAPKRQGFCTIFQLVSTESNRY LDHISLFTSQETLIDFALTSTDIWALHDAENQTVVKYINFEHNVAGQWNPVFMQPL PEEETIVRDDQDPREMYLQSLFTPGQFTNEALCKALQIFCRGTERNLDSLWSELKKEV TLAVENELQGSVTEYEFSSQEEFRNLQEEFWCKFYACCLQYQOALSHPLALHLNPHNM VCLLKGYSFLIPSSLVHDHLYLLPYENLLTEDETTISDDVDIARDVICLIKCLRLIE ESVTVDMSVIMEMSCYNLQSPKAAEQILEDMITIDVENVMEDICSKLQEIIRNP GLLIREMDYETEVEMEKGFNPAQPLNIRNMLTQLYGSENTAGYIVCRGVHKI CRDLLILQQLMLRLGDAVIWGTGQLFQAQDQLLHRTAPLLLSYLLIKWSECLATDVP LDTLESLNLQHLVLELTDGALMANRFVSSPQTIVELFFQEVARKHIIISHLFSQPKAP LSQTGLNWPMTITATSYLLQLLWPSNPGCLFLECLMGNCQYVQLQDYIQLLHPWCQV NVGSCRFMLGRCYLVTGEGQKALECFQAAAEVKGEEFLDRILIRSEDGEIVSTPRLOY YDKVLRLDLVIGLPELVIQLATSAITEASDDWKSQATLRTCTFKHHLDLGHNSQAYEA LTQIPDSSRQLDCLRLVVLVLCERSQLQDLVEFPYVNLHNEVVGIIESRARAVDLMT NYIELLYAFHIYRHNRYRKAGTVMFEYGMRLGREVRTLRGLEKQGNCYLAALNCLRLIR PEYAWIVQPVSGAVYDRPGASPKRNHDGECTAAPTNRQIEILELEDLEKECSLARIRL TLAQHDPASAVAVAGSSSAEEMVLLVQAGLFDTAISLCQTFKPLPLTPVFEGLAFKCIK LQFGGEAAQAEAWAWLANQLSSVITTKESSATDEAWRLSTYLERKYVQNLYKHCV INKLLSHGVPLPNWLINSYKKVDAEALLRLVLYDLLEEAVDLVSEYDAVLGKHGQY FGIEVGIYKWSPIGRDSANSHNIALSQKILDKLEDYQKVKDKATRDLLYRRTL
	SEQ ID NO: 179 5264 bp
NOV61b, CG98061-02 DNA Sequence	GCTCTTCTGGATTCTCGCTGCCACCTCATTCTCCCGTGGAAGCCTCGCAACGAA GGGCCGGAGCCGCTTTCTGCTCCCGGAATGCTTACCTGTCGCGAGCTCCGCCCGCC CCACCCCGGAAGTGACGGCGACCGCGCGGCTGCTTTGTTCGTTGGGCGTCCGCGG CGACGGCGGGAAGATGGCGGCGGCGGAGCCCTGGAACGGAGCTTCGTGGAGCTAAGC GGAGCTGAGCGCGAAAGCCGAGGCCTTTCCGGAATTCACAGTCTGCAGCATTTGGGA CTGCAATGCCGTGGCTGGCGCGGTAAATACAGTGAAAGCGCGGGAGGCTTTTACTA CGTGGAGAGTGGCAAGTTGTTCTCCGTAACAGAAACAGGTTTATTCTTGGAGAGCC TCTGGAGATACATTGGAGCTGATGGAGGAGTCACTGGACATAAATCTGTTGAATAATG CCATTCCGCTAAATTCAAAATTCAGTGTGTTTACCTGGAGGGTTTATGTTCTCTGA GACTCAGAACTCGTGTGATAATCTTGATGTTAACCAATCAACAGTGCACAGGTTACTT TTACCACACCCCTCCCGGATGTATAGGAGTGAGTTGGTAGTTGACAGTCAGATGCAGT CAATATTCACTGACATTGGAAGAGTTGATTTCACAGATCCTTGCACTATCAGTTAAT TCCAGCAGTACCTGGAATATCTCCTAATCCACCGCTCTACAGCCTGGCTCAGCAGT GATGGGGAGGCGCTGTTGCTTACCATGTGCTTCTGGGGGAATCTTTGTTCTTAAGC TACCTCCTTATGACATACCTGGTATGGTGTGAGTCTGCTGGAAGTGAACAGAGTTCA AATGCAACGATTGCTTACAGGCTGGATGCCAACAGCTATCAGGGGTGACCACTCGCCT TCAGATCGTCCCTCAGTCTTGCTGTTTATGTTGGAGCATGATGCCTTATCTTTG CTTTGTGTGAGGATCATAAATACGAATGTGCTTACAAGGAGCAATGTGCCTAAT GGTAGCTGACATGCTGGAGTATGCTCCTGTGAAGAAAGACCTTCGGCTTACTGCTGGA ACTGGACACAAATTACGGCTTGCTTATCCCCCACCATTGGGACTCTACCTGGGGATAT ACATGCATGCACAAACAGGAGCAGTTCTGCATTTTCCAGTTGGTGAGCACTGAGAG TAATCGTATAGTCTCGATCATATTTCTTCACTGTTCACTTCTCAGGAGACACTGATT GACTTTGCTTAACTTCCACGGATATCTGGGCCCTGTGGCATGATGCTGAGAACCAAA CAGTAGTGAAATACATCAACTTTGAACATAATGTTGCAGGTGAGTGAATCCAGTTT TATGCAGCCTCTGCCAGAGGAAGAGATTGTCATCAGAGATGATCAAGACCCAGAGAG ATGTATCTGCAAGTCTTTTACACAGGACAATTCACAAATGAAGCTTTATGTAAGG CTTTACAGATTTCTGCCGAGGAAGTGAAGGAATTTGGATCTTTCTGGAGTGAAT GAAGAAAGAGTTACTTTAGCTGTTGAAATGAGCTTCAAGGAAGTGAACAGAGTAT GAATTTCTCCAGGAGGAGTTTCGAAATTTACAACAAGAAATTCGGTGCAAGTTCTATG CCTGTTGTCTTCAATATCAAGAACCCCTCTCACCCTCTTGCCCTACATTTGAATCC ACACACAAACATGGTGTGCTGCTGAAAAAGGGTACCTGTCTTTCTTATCCCTCA TCCTTAGTGGATCATTTGTATCTCCTGCTTATGAGAACCTTTTGACAGAGATGAGA CAACCATATCTGATGATGTGGATATCGCTCGGGATGTCATATGCTTTATAAATGCTC CCGGCTGATTGAAGAGTCAGTAACGTGGATATGTCAGTTATAATGGAATGAGTTGT TATAACCTACAGTCTCCGGAAGGGCTGCAGAGCAGATTCTGGAAGATATGATCACTA TTGATGTAGAAATGTGATGGAGGATTTTGTAGTAACTGCAAGAGATTAGGAACCC AATCCATGCAATTGGACTTATACCGGAATGGATTATGAACCAAGTGGAAATG GAAAGGGATTCAATCCAGCTCAGCCTTTGAATATTGCAATGAATCTTACCCAGCTCT

ATGGTAGTAACACAGCAGGGTATATTGTGTGCAGAGGGGTGCATAAAATCGCCAGTAC
TCGTTTCCTGATCTGCAGAGATCTTTTGATCTTACAGCAGCTGTTAATGAGGCTTGGA
GATGCTGTGATTTGGGGAACTGGTCAGCTCTTCAAGCTCAGCAAGACCTACTACATC
GAACAGCTCCCCACTCTTATCTTATTACCTCATTAAATGGGGAAGTGAGTGTCTGGC
AACTGATGTTCCACTTGACACACTGGAGTCTAATCTCCAACACTTATCAGTACTGGAA
TTAACAGACTCTGGTGCCTTAATGGCAAATAGGTTTGTATCTAGTCCCTCAGACTATTG
TGGAGTTATTCTTCCAAGAAAGTTGCAAGAAAACACATTATATCTCACCTCTTCTCTCA
GCCAAAGGCACCTCTGAGCCAAACTGGATTGAATTGGCCTGAAATGATTACTGCAATT
ACCACTTATTATTGTCAGCTTTTATGGCCTAGCAATCCTGGTTGTCTCTTCTAGAAAT
GTTTGATGGGAAATTGCCAATATGTACAATTGCAGGATTATATTCACTGCTACATCC
CTGGTGTCAAGTCAATGTTGGTCTGTCTGATTTATGCTGGGAAGGTGTACCTAGTT
ACAGGAGAAGGACAGAAGGCTCTGGAATGTTTGTGTCAGGCAGCATCTGAAGTAGGCA
AAGAGGAATTCTTGGATCGCTTGATTGCTCAGAGGATGGGGAGATCGTGTCTACCCG
CAGGCTGCAGTATTATGACAAGGTTTACGACTACTAGATGTCAATGGTTTGGCTGAA
CTGGTTATTAGTTGGCTACATCAGCCATAACTGAAGCAAGTGATGACTGGAAAAGTC
AGGCTACTCTAAGGACATGTATTTCAAACATCATTGGATTGGGTCACAATAGGCCA
AGCATATGAAGCCTTAACCCAAATTCCTGATTCAGCAGGCAATTAGATTGTTTACGG
CAGTTGGTGGTATTCTTGTGAACGCTCAGAGTACAGGATCTTGTAGAGTTTCCCT
ATGTGAATCTGCATAATGAGGTTGTGGGAATAATTGAGTCAGTGTCTAGAGCTGTGGA
CCTTATGACTCACAATTACTATGAACCTTCTGTATGCTTTTACATCTATCGCCACAAT
TACCGCAAGGCTGGCAGTGATGTTTGAGTATGGAATGCGGCTTGGCAGAGAAGTTC
GAACTCTCGGGGACTTGAGAAACAAGGCAACTGTTATCTGGCTGCTCTCAATTGTTT
ACGACTTATTCGTCAGAAATATGCGTGGATTGTGTCAGCCAGTGTCTGGTGCAGTGTAT
GATCGCCCTGGAGCATCCCCAAGAGGAATCATGATGGAGAATGCACAGCTGCCCCCA
CAAAATCGACAAATGAAATCCTGGAACCTGGAAGATCTGGAGAAAAGTGTCTTCTGGC
TCGCATCCGCTCACTTTGGCTCAGCATGATCCATCAGCGGTTGCAGTGTCTGGAAGT
TCATCAGCAGAGGAAATGGTCACTCTCTTGGTTCAGGCGGCTCTTTGACACTGCCA
TATCACTCTGTGAGCTTTAAGCTTCCCTTAACGCCAGTCTTTGAAGGGCTTGCCCTT
CAAAATGCATCAAAATGCAATTTGGAGGAGAGGCAGCACAAGCAGAGCTTGGGCTGG
CTAGCAGCCAAATCAGCTCTCATCTGTCTACTACTAAGGAGTCTAGTGTACAGATG
AAGCATGGCGACTATTATCCACTTACCTGGAGAGGTACAAGTCCAGAATAACTTGTA
TCACCACTGTGTAAATCAACAAGCTCTGTCTCATGGAGTGCCTCTGCTCAATTGGCTT
ATAAACAGTTACAAGAAGGTTGATGCTGCTGAATTGCTTCGTTTATACTTAACTATG
ACCTTTTAGAAGAAGCTGTGGATTGGTGTGAGAATATGTTGATGCTGATTGGGAAA
AGGACATCAATACTTCGGAATTGAGTTTCCACTGTCCGCAACAGCCCCAATGGTGTGG
CTTCCATACTCCTCTATTGATCAGCTTCTCCAAGCTCTGGGAGAGAAGCTTGGGCTGG
GTCACAACATCGCACTGTCCAGAAAATACTTGACAAATGGAGGACTACCAGCAAAA
AGTTGATAAGGCAACACGGGATTATTATATCGTCGGACCTTGTGATTGGATTGTCA
CCTAGCCTTTGTAACCGCTTGGTGCCTCTTAGGACTTAAGACTACCTACAGGAACCC
TGTAATCAAGGCCGATTTTGTAACTGTAATGATGTGTACAACATTCAAGTCTGCAT
TCTGCACAAGATAGGAGGGCGGAAGAGTCAGAGGACCCTGTGCTTGGTGGTGTGCTA
ACACAATTTCTGGTGTCAACCTTGGTCTCAAAATAGCTGCTTTTGTATATGATTACG
AGCTTTTTTAGAGTTTATATTTTTTAACTACCGAAGACATTCTATCTGCAAAAT
AAGACTCACCTTCACTTTCCAAAATAGCTGAGGGTTGTTGGCTTGTGTAGCTGACCA
CCAAAAGCAGTCACTGCAAAATCTTTAATTCTTCCCTATCACCTTTGTATTTAATG
CAATTATTTGGTCCAGAACTGACCTGTATTTCTGTATTGTACACAAAAGCTAATAA
TTTTGTGTACTTTTATTATTTTGGAGGTTTATATGATCTTCAATTGAGTATTAAA
TAATTGGCTAGATTAAGCCTAAAATGATGACCAGCTAATTAAAGAAGATATTTGAA
TCTGGTCTGAGCTAAAGTTGAGTAAATCTTAGCTAAGAAAAAATGGAAATCCATC
ATCTATATTAGCAACAGATTCTCAGAGTAAATGTTAACTTCTATGATTATGATAAT
CAAGCTGGACTTGATCATAAAGTTAGTCTCATAATGTATTGGACCAAAATGTAACT
TCATTGGTCAGATTTAGAAGCATTATGCTCACAAGTTTTGGGAAAGTGAAAAATAAT
AAAATCATCTGGATTATTCTGTATATTTAAATTTATCTTT

	ORF Start: ATG at 188	ORF Stop: TGA at 4394	
	SEQ ID NO: 180	1402 aa	MW at 158753.4kD
NOV61b, CG98061-02 Protein Sequence	MAAGALERSFVELSGAERERPRHFREFTVCSIGTANAVAGAVKYESAGGFYYVESG KLFSVTRNRFIHWKTSKDTLELMEESLDINLLNNAIRLKFQNCVSLPGGVYVSETQNR VIIIMLTNQTVHRLLLPHPSRMYRSELVVDSDMQSIFTDIGKVDFTDPCNYQLIPAVP GISPNSTASTAWLSSDGEALFALPCASGGIFVLKLPYDIPGMVSVVELKQSSVMQRL LTGWMPTAIRGDQSPSDRPLSLAVHCVCHDAFIFALCQDHKLRMWSYKEQMCLMVDAM LEVVPVKKDLRLTAGTGHKLRLAYSPTMGLYLGIMHAPKRGQFCIFQLVSTESNRYSL LDHISSLFTSQETLIDFALTSTDIWALWHAENQTVVKYINFEHNVAGQWNPVFMQPL PEEEIVIRDDQDPREMYLQSLFTPGQFTNEALCKALQIFCRGTERNLDSLWSSELKKEV TLAVENELQGSVTEYEFSSQEEPRNLQOEFWCKFYACCLQYQEALSHPLALHLPHTNM VCLLKKGYLSFLIPSSSLVDHLVLLPYENLLTEDETTISDDVDIARDVICLIKCLRLIE ESVTVDMSVIMEMSCYNLQSPKAAEQILEDMITIDVENVMEDICSLKQEIIRNPIHAI GLLIEMDYETEVEMEKGFNPAQPLNIRMNLTQLYGSNTAGYIVCRGVHKKIASTRFLI CRDLLILQQLLMRLGDVIVGTGQLFQAQDQLLHRTAPLLSYLIKWSECLATDVP LDTLESNLQHLVLELTDGALMANRPFVSSPQTIVELFFQEVARKHIIISLFSQPKAP LSQTGLNWPemitaitSYLLQLLWPSNPGCLFLECLMGNCQYVQLQDYIQLLHPWCQV NVGSCRFLMGRCYLVTGEGQKALECFQAASEVKGEEFLDLIRSEDGEIVSTPRLQY YDKVLRLLDVIGLPELVIQLATSAITEASDDWKSQATLRTICIFKHHLDLGHNSQAYEA LTQIPDSSRQLDCLRQLVVLCCERSQLQDLVEFPYVNLHNEVVGIIESRARAVDLMT NYEYELLYAFHIYRHNRYKAGTVMFEYGMRLGREVRTLRGLEKQGNCYLAALNCLRLIR PEYAWIVQPVSGAVYDRPGASPKRNHDGECTAAPTNRQIEILELEDLEKECSLARIRL TLAQHDPSAVAVAGSSAEEMVTLVQAGLFDTAISLCQTFKLPLTPVFEGLAFKCIK LQFGGEAAQAEAWLAANQLSSVITTKESSATDEAWRLSTYLERVKVQNNLYHHCV INKLLSHGVPLPNWLINSYKKVDAEELLRLYLNYDLLEEAVDLVSEYVDAVLGKGHQY FGIEFFPLSATAPMVWLPYSSIDQLLQALGENSANSNIALSQIKLDKLEDYQQKVDKA TRDLLYRRTL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 61B.

Table 61B. Comparison of NOV61a against NOV61b.		
Protein Sequence	NOV61a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV61b	1..1387 1..1402	1308/1402 (93%) 1312/1402 (93%)

Further analysis of the NOV61a protein yielded the following properties shown in
5 Table 61C.

Table 61C. Protein Sequence Properties NOV61a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV61a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 61D.

Table 61D. Geneseq Results for NOV61a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV61a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAM40894	Human polypeptide SEQ ID NO 5825 - Homo sapiens, 1401 aa. [WO200153312-A1, 26-JUL-2001]	1..1314 75..1389	1308/1315 (99%) 1311/1315 (99%)	0.0
AAM39108	Human polypeptide SEQ ID NO 2253 - Homo sapiens, 1316 aa. [WO200153312-A1, 26-JUL-2001]	1..1307 1..1312	1299/1312 (99%) 1301/1312 (99%)	0.0
ABB60919	Drosophila melanogaster polypeptide SEQ ID NO 9549 - Drosophila melanogaster, 1411 aa. [WO200171042-A2, 27-SEP-2001]	43..1335 41..1333	389/1342 (28%) 638/1342 (46%)	e-143
AAB38261	Gene 13 human secreted protein homologous amino acid sequence #117 - Homo sapiens, 58 aa. [WO200058469-A1, 05-OCT-2000]	471..528 1..58	58/58 (100%) 58/58 (100%)	9e-29
AAB38262	Human secreted protein sequence encoded by gene 13 SEQ ID NO:118 - Homo sapiens, 58 aa. [WO200058469-A1, 05-OCT-2000]	471..528 1..58	57/58 (98%) 58/58 (99%)	2e-28

In a BLAST search of public sequence databases, the NOV61a protein was found to have homology to the proteins shown in the BLASTP data in Table 61E.

Table 61E. Public BLASTP Results for NOV61a				
Protein Accession Number	Protein/Organism/Length	NOV61a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q12769	Protein KIAA0197 - Homo sapiens (Human), 1314 aa (fragment).	3..1307 1..1310	1298/1310 (99%) 1300/1310 (99%)	0.0
Q9Z0W3	Protein KIAA0197 (GTL-13) - Mus musculus (Mouse), 1402 aa.	1..1387 1..1402	1274/1402 (90%) 1332/1402 (94%)	0.0
Q9VKJ3	Protein KIAA0197 homolog - Drosophila melanogaster (Fruit fly), 1411 aa.	43..1335 41..1333	389/1342 (28%) 638/1342 (46%)	e-143
Q9CZD9	2810011M03RIK PROTEIN - Mus musculus (Mouse), 215 aa.	1181..1380 1..215	180/215 (83%) 187/215 (86%)	1e-96
Q96GB3	KIAA0197 PROTEIN - Homo sapiens (Human), 190 aa.	1..141 1..141	141/141 (100%) 141/141 (100%)	5e-75

PFam analysis predicts that the NOV61a protein contains the domains shown in the Table 61F.

Table 61F. Domain Analysis of NOV61a			
Pfam Domain	NOV61a Match Region	Identities/ Similarities for the Matched Region	Expect Value
No Significant Matches Found			

Example 62.

- 5 The NOV62 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 62A.

Table 62A. NOV62 Sequence Analysis

	SEQ ID NO: 181	5544 bp
NOV62a, CG98071-01 DNA Sequence	<p>GCGGGACTCAAGAGTAGCCTTCCTCGAGGACCTGCCTTTCCTCATTTGCTGCCTGAAGT TAATGTTTCTTGTGCTGGCCAAATCAGGGACATGCCGGCATTAGCGGGATGAGTGGGTGT TCCGGCAGGGATGTGGTCTATTGACGGCCAGTGAGGGCGAGAGTACCACGGCCCCACTTC TTCCTTGGAGCTGGAGATGAGGGGCTGGGCACCCGTGGAATAGGCATGAGGCCAGAAG AGAGTGACAGCGAGCTCCTTGAGGATGAGGAGGATGAAGTGCTCCTGAACCTCAGAT CATTGTTGGCATCTGTGCCATGACCAAGAAATCCAAGTCCAAGCCAATGACTCAATC CTAGAGCGACTCTGCAGATTGACTACCTGACTGTTGTCATTCTGGGAGAAGATGTAA TCCTTAATGAACCTGTGGAAAATGGCCATCCTGCCACTGCCTCATCTCTTCCACTC CAAAGGCTTTCCTGGAACAAGCTGTTGCTTACTCCAAGCTTCGAAACCCCTTCTT ATCAATGATCTGGCCATGCAGTATTACATCCAAGATAGGAGGGAGGTGTACCGGATCC TGCAGGAAGAGGGTATTGATCTGCCTCGATATGCTGTGCTCAACCGTGATCCTGCCCG GCCTGAGGAATGCAACCTGATAGAAGGTGAAGACCAAGTAGAGGTCAATGGAGCTGTC TTTCCCAAGCCCTTGTGGAGAAGCCAGTGAGTGAGAGGAGGAGGTGAACGAGACAGTGA ACTACCCAGCTCAGCTGGAGGAGGAAGCCAGCGTCTCTTTCGTAAGATTGGCAGCCG AAGCAGTGTTTACTCTCCTGAGAGCAGCGTCCGAAAGACGGGGTCGTACATCTATGAG GAGTTTATGCCAACAGATGGCACAGATGTCAAGGTGTATACAGTGGGGCCAGATTATG CCCATGCTGAAGCTAGAAAATCTCCAGCTTGGATGGGAAGGTGAACGAGACAGTGA GGGGAAGAGATTGATATCCAGTCATGCTGACTGCCATGGAAAAGCTGGTGGCCAGG AAAGTCTGCGTAGCTTCAAGCAACAGTGTGTTGGATTGACCTTCTCGTGCCCAATG GTCATTCTTGTGTGTGATGTCAATGGCTTTAGTTTGTCAAGAACTCGATGAATA CTACGATGACTGTGCCAAGATTCTGGGAACACCAATAATGCGGGAGCTTGGCCACAG TTCCAGATTCCATGGTCCATCCCAACGGAGGCTGAGGACATTCCCATTTGTTCCACCA CATCTGGCACTATGATGGAATTCGTTGTGTCATTGCAATTATTGTCATGGGGATCG TACTCCCAAGCAGAAGATGAAGATGGAAGTGAACACCCCAAGCTTTTGTCTGTGTT GAAAAACATGGTGGCTACAAGACAGGGAATTAAGCTCAAGCGACCTGAGCAGCTCC AGGAGGTGCTGGATATCAAGGCTGTTGTTGGCTGAATGGAGAAAGAACAGGTGG TGAGATCGAGGAGAAGACTGGAAAATAGAGCAGCTGAAGTCTGTACTGGAGATGTAT GGTCATTCTCAGGTATAAACCGGAAGGTACAATTGACTTACTACCTCATGGAGTAA AAGCTTCTAATGAGGGGCAAGATCCACAGAGGGAAGCTGCGCCCCATCTCTGTTGCT GGTACTGAAGTGGGGTGGAGAAGTACTCCTGCTGGCCGTGTTCAAGCTGAGGAGCTG GGGCGAGCTTTTCTGTCATGTACCTGAGGACAGGGTGAATGCTGGCTTCCCTG GTTGTGGGCTGCTTCTCTCCATAGCACTTTCGCCACAGATCTCAAGATCTATGCCTC TGATGAGGCTGCTGTTGATGACTGCTGCTGCTTCCGCAAGGGCTTCTGGCTCTA GAAGGGAGCTGACACCATTTTGGTGCAATGGTGAAGAGTGCCAAATGAATGGGC TACTGGACAGCGATGGGATTCCTTGAGCAGCTGCCAGCACCGGTTGAAGCTCGGCT GCACCATATTCTACAGCAGGATGCACCTTTGGCCCTGAGGACTACGATCAGCTGGCT CCCACCAAGTACTTCCCTGCTCACTCCATGACTATCATCCAGAATCCTGTGAAGG TCTGTATCAGGTATTGCCCCGCTGATCGAAAACCTCACCCACAGATCCGGGAACGAAT GCAGGACCCAGGTCTGTAGACCTGCAGCTCTACCAAGTGAGACACTAGAGCTAATG CTACAGCGTTGGAGCAAGCTGGAGCGTGACTTTGACAGAAGAGTGGGCGCTATGATA TCAGTAAGATCCCTGACATCTATGACTGTGTCAAGTATGATGTCAGACAAATGGGAG TCTGGGACTTCAAGGCACAGCAGAGTTGCTCCGCTCTCTAAGGCACTGGCTGATGTG GTCATTCCCCAGGAGTACGGGATCAGTCGGGAGGAGAACTGGAAATGCTGTGGGCT TCTGTCTTCCACTGTTGCGGAAGATACTACTTGACCTGCAGAGAACCACAGGAGTGA GTCTGTCAACAAGCTGCATCCCTGTACTCCGAGGCGTGCTCTCCCGAGTCCGCCAC GTTGCAAGCGCTCTTATTTACCAAGTGAGAGCCATGTCCACTCCCTGCTCAGTGTCT TCCGTTATGGAGGACTTCTGATGAGACCCAGGATGCACAATGGCAGCGAGCTTTGGA TTATCTTAGTGCCATCTCAGAGCTTAATACATGACCCAGATTGTCTATCATGCTTTAT GAGGACAACACACAGGATCCCTTATCAGAGGAACGGTTCCATGTGGAGCTACACTTCA GCCCCGAGTGAAAGGTGTTGAGGAAGAAGGCAGTGCCCGGCTGGCTGTGGATTCCG TCCAGCCTCTTCTGAGAATGAGGAGATGAAAACCAACCAAGGCAGTATGGAGAACCTG TGTCCAGGAAGGCATCAGATGAACAGACCGGGCATTCAGACTTCACCCAGCCTC CTGAGGCGCTTGGCTTCCGAGGAGATCACCCCTCATTGTAACCGAAAGCTGGTTTC CATGGAGGTACTTTCTGAGACTTCATCCTCGAGGCTGGTGGCTACCGGCTCTTTTCA TCTTACAGGCCACCCACAGAAATGAAGCAGAGTGGCTTAGGGTTTGAAGGTTGTTCCA TGGTGCTACCATCTACCCTCTGGAACACTGCATAATGCCCTTTCCCTACGTCAGT GAGTGAATCTTGAGTAGAGTCTGCCAGCGCCACACTGATGCCAGGCACAGGCATCT GCAGCCCTCTTTGATTCCATGCACAGCAGCCAGCCCTCAGATAACCCATTTTCTCCAC CAGCTACTCTTCACTCACCTCCCTGCAACTCCAGCAGCGCTCTGAGAAGCCCCCTTG GTTAGAGACAAGGTTTGGCATGTTGGCCAGGCTGGTTTAGAGCTCTGACCTCAAGT GATCTGCTGCTCGGCTCCCAAGTGTGGGATTACAGGCGTGAGCCACCCGACCC AGCCAGACAGCAGTGGCCCTTCTAGCACTGTGTCCAGTGCTGGTCTCTTCTCCCTAC TACAGTAGATGTAATCCCAATTGGCTTCAAGTATCAACCTCCCTAAATTCACAC GTGGCTGAAGAACATCAAGCCCTTGGGCTGCTCCAGGAGACCCCTGGGAGTGGAGCAC AAGAGCTCTCCATAGAAGGGGAGCAAGAGCTTTTGAACCAATCAGTCCCCACAGGT GCCACCTATGGAACACGACGACCATACGAGGAGGTGAGCCAGCCATGTGAGGAGGTC CCTGACATCAGCCAGCCATGCCAGGACATTTCTGAGGCGCTCAGCCAGCCATGTGAGA</p>	

	<p>AGGTCCTTGACATCAGCCAGCAATGCCAGGAGAACCATGCAATGGTAACCACACATG CCAGGAGGTCCCTCACATCAGCCAGCCATGCCAGAAGTCCAGCCAACTGTGCCAGAAA GTCTCTGAGGAAGTTTGCCAGCTATGTCTGGAGAACTCCGAGGAGGTGAGCCAGCCAT GCCAGGGGGTCTCTGTGGAGGTTGGCAAGCTGGTCCATAAGTTCATGTAGGGGTTGG TAGCTTGGTCCAGGAAACCCTTGTAGAAGTTGGCAGCCAGCTGAAGAGATCCCTGAG GAGGTCATCCAGCCATACCAGGAGTTCTCTGTGGAGGTTGGCAGGCTGGCCAGGAGA CTTCTGCGATCAATCTGTTATCTCAGGGCATCCCTGAGATTGATAAACCATCCAGGA GTTCCCTGAGGAGATTGATCTGCAGGCCAGGAGGTCCCTGAGGAGATAAATTAGAAAG TCCTGGGTGGTCCCTGAAGTGATTGATCAGCTGCCTGGAGAGGGTATTCTCAAGCCC AGCATCCATCTGGTGATCCAAACCCTCAGAGCCAGTCTCTAGCCCATGACCAGCACTC ACCCCTTCCACCAGCAACATGTGATTAAATTTCTCATTAGTGGTATCACACTATACCA GCCATTTGAGCCAGCAACCTTTTCTGTTGGCTAACTCACTGGCCAGCTCTACCAGCG GTGTCTGGGGAAGTAGTTCTCTTGTATGAAGCATACCTGTGCCAGAGCTGTGGTTGA GGAGGAGCCAGTTTATAGTTTGAAGAAGCCATTGGCTCTCACTTAGCCATTAGACTT GAATAGGATTTCTTGGGGTGGGTTGGTCTGTCTATACCAGCCCTTCTCTGAGCATCC TAGGAAATCACAGATTGTTAAAGGAAATGCCGCTTCACTGCTGAAGACACCATCTGGC GACAGCAATGCAAAAGAGGGGACTCTAGGGTCTTCACTTTCTGGGAAATGTTTAC GACTTCTCAAGGTACGCTTAGACCATATGTGCATTGAGGGGACTTGTCTTTGCCAG CACTCCTCAGGTGAGCCCGGATGGCTATCTGAAATGTCTGGAATAATAGAGTCCCTTC CCCAAACATCTGACATGGCCACTAAATCTTTAGACTTGGCTTATAGATCCTTCCATAT CATCCCCAAGTGGCCCTTCTGCTCAGTCTCTGTTCTCTGGGCAATAGCTCTAGGGAA AAGTAGGTATTAGACTGCTGTGCAAAATTCAGAGCAACTACTTAAATGCTCTAGAGG ATTCTTAGCCAGTCTGTGAAAGTGGGCTCCTCTGTGGTGGGGCTGTTTGGCTTAGGA GATGCTGTAGTAGTAGAGATAATCAGGTTCTATTTTAAAGCAATCAGCAGAGATCAATA TTGTACAGATACAAGCAAAGGTTTAAATAATATATATATATATTTATTCTGTAGT TGTGCCAGGACAGACTGGCCAAAGACCAACTGCAGGGTCCCCAAGAAATTTGTCT CTTATATCATTGTGTCTGAGGGCCAGATATGATTATGAAGCTTTTCCAAAGATCCAG GGATGGGAATGGGAGTGGGTAGGAGGGGTAAATGCGGTCTTGGAGTCTGGGGGCTGGAA CATTATGAGTGCTCAATAAATAAATAAATAATGAG</p>
	<div>ORF Start: ATG at 127</div> <div>ORF Stop: TAG at 4345</div>
	<div>SEQ ID NO: 182</div> <div>1406 aa</div> <div>MW at 156318.6kD</div>
NOV62a, CG98071-01 Protein Sequence	<p>MWSLTASEGESTTAHFFLAGDEGLTRGIGMRPEESDSELLEDEEDEVPEPQIIVG ICAMTKKSKSPMTQILRLCRFDYLTVVILGEDVILNEPVENWPSCHLISFHSKGF PLDKAVAYSKLRLNPFLLINDLAMQYIYQDRREVYRILQEEGIDLPYAVLNRDPARPEE CNLIEGEDQVEVNGAVFPKPFVEKPVSAEDHNVIYYPSSAGGGSQRLFRKIGSRSSV YSPSSSVRKTSYIYEEFMPDTGTDVKVYTVGPDYAHAEARKSPALDGKVERDSEGKE IRYPVMLTAMEKLVARVVCVAFKQTVCGFDLLRANGHSFVCDVNGFSFVKNSMKYDD CAKILGNTIMRELAPQFIWPSIPTAEEDIPVPTTSGTMMLRCVIAIIRHGDRTPK QKMKMEVKHPRFFALFEKHGGYKTKLKLKRPEQLQEVLDITRLLLAELKEPGEIE EKTGKLEQLKSLVLEMYGHFSGINRKVQLTYYPHGVKASNEGQDPQRETLAPSLLLVLK WGGELTPAGRVQAEELGRAFCMYPGGQGDYAGFPCCGLRLHSTFRHDLKIYASDEG RVQMTAAAFAGLLALEGELTPILVMVKANMNGLLDSDGDSLSSCQHRVKARLHHI LQGDAPFGPEDYDQLAPTRSTSLNSMTIIQNPVKVCDQVFALINLTHQIRERMQDP RSVDLQLYHSETLEMLQWRWSKLERDFRQKSGRYDISKIPDIYDCVKYDVQHNGSLGL QGTAEALLRLSKALADVVIPOEYGISREEKLEIAGVFCPLLRKILLDLQRTHEDESVN KLHPLYSRGVLSPGRHVRTRLYFTSESHVHLLSVFRYGLLDETDQAQWRALDYLS AISELNMTQIVIMLYEDNTQDPLSEERFHVLFHSPGVKGVEEGSAAPAGCGFRPAS SENEEMKTNQGSMENLCPGKASDEPDRLQTSQPPEGPGLPRRSPILNRKAGSMEV LSETSSSRPGGYRLFSSSRPPTMKQSGLGFEGCSMVPTIYPLETLHNLALSLRQVSEF LSRVCQRHTDAQASAAALFDSMHSSQASDNPFSPRTLHSPPLQLQQRSEKPPWLET RFCHVGQAGLELLTSSDLPASASQASAGITGVSHRTQPDSSGSPSTVSSAGPSPPTVD GNSQFGFSDQPSLNSHVAEEHQGLGLLQETPGSGAQELSTIEGEQELFEPNQSPQVPPM ETSQPYEEVSQPCQEVDPDISQPCQDISEALSQPCQKVPDISQCCQENHDNGNHTCQEV PHISQPCQKSSQLCQKVSEEVQCLLENSEEVSPCQGVSEVKGKLVHFKFHVGVGSLV QETLVEVGSPAEIPEEVIQPYQEFSEVVGRLAQETSAINLLSQGIPEIDKPSQEPPE EIDLQAQEVPEIN</p>

Further analysis of the NOV62a protein yielded the following properties shown in Table 62B.

Table 62B. Protein Sequence Properties NOV62a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3000 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV62a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 62C.

Table 62C. Geneseq Results for NOV62a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV62a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG15880	Novel human diagnostic protein #15871 - Homo sapiens, 1136 aa. [WO200175067-A2, 11-OCT-2001]	161..1008 239..1056	621/850 (73%) 695/850 (81%)	0.0
ABG14341	Novel human diagnostic protein #14332 - Homo sapiens, 1136 aa. [WO200175067-A2, 11-OCT-2001]	161..1008 239..1056	621/850 (73%) 695/850 (81%)	0.0
ABG15880	Novel human diagnostic protein #15871 - Homo sapiens, 1136 aa. [WO200175067-A2, 11-OCT-2001]	161..1008 239..1056	621/850 (73%) 695/850 (81%)	0.0
ABG14341	Novel human diagnostic protein #14332 - Homo sapiens, 1136 aa. [WO200175067-A2, 11-OCT-2001]	161..1008 239..1056	621/850 (73%) 695/850 (81%)	0.0
ABB69600	Drosophila melanogaster polypeptide SEQ ID NO 35592 - Drosophila melanogaster, 1751 aa. [WO200171042-A2, 27-SEP-2001]	16..954 33..991	619/966 (64%) 730/966 (75%)	0.0

- 5 In a BLAST search of public sequence databases, the NOV62a protein was found to have homology to the proteins shown in the BLASTP data in Table 62D.

Table 62D. Public BLASTP Results for NOV62a				
Protein Accession Number	Protein/Organism/Length	NOV62a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
O15082	KIAA0377 PROTEIN - Homo sapiens (Human), 1406 aa.	1..1406 1..1406	1406/1406 (100%) 1406/1406 (100%)	0.0
AAH24591	SIMILAR TO KIAA0433 PROTEIN - Homo sapiens (Human), 1222 aa.	11..997 2..986	743/989 (75%) 835/989 (84%)	0.0
O43314	KIAA0433 PROTEIN - Homo sapiens (Human), 1243 aa.	11..997 2..986	743/989 (75%) 835/989 (84%)	0.0
Q9VR59	CG14616 PROTEIN - Drosophila melanogaster (Fruit fly), 1751 aa.	16..954 33..991	619/966 (64%) 730/966 (75%)	0.0
T25770	hypothetical protein F46F11.1 - Caenorhabditis elegans, 1224 aa.	54..909 16..917	491/910 (53%) 627/910 (67%)	0.0

PFam analysis predicts that the NOV62a protein contains the domains shown in the Table 62E.

Table 62E. Domain Analysis of NOV62a			
Pfam Domain	NOV62a Match Region	Identities/ Similarities for the Matched Region	Expect Value
acid_phosphat	488..947	78/531 (15%) 290/531 (55%)	0.4
zf-NF-X1	1269..1292	8/27 (30%) 20/27 (74%)	0.75

Example 63.

- 5 The NOV63 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 63A.

Table 63A. NOV63 Sequence Analysis			
	SEQ ID NO: 183	885 bp	
NOV63a, CG98131-01 DNA Sequence	CGGCCACTTCTCCCCCGTCCATCTGACCAGCTGGGCCCTGCGCCACCTGGCCTCCA CGTTCCTCTCCTCTCACCCACACCCCTGGCCATGGCTAACTACTATGAAGTGCTGGG CGTGACAGCCAGCGCTTCCCCGGAGGACATCAAGAAAGCCTACCGCAAGCTGGCCCTT CGTTGGCACCCGACAAGAACCCTGACAATAAGGAGGAGGCGGAGAAGAAGTTCAAGC TGGTGTCTGAGGCCTATGAGGTTCTGTCTGACTCCAAGAAACGCTCCCTGTATGACCG TGCTGGCTGTGACAGCTGGCGGGCTGGTGGCGGGGCCAGCAGCCCTACCACAGCCCC TTCGACACCGGCTACACCTTCCGTAACCCTGAGGACATCTTCGGGAGTTTTTTGGTG GCCTGGACCCCTTCTCCTTTGAGTTCTGGGACAGCCATTCAATAGTGACCGTGGTGG CCGGGGCCATGGCCTGAGGGGGGCTTCTCGGCAGGCTTTGGAGAATTTCCGGCCTTC ATGGAGGCCTTCTCATCCTTCAACATGCTGGGCTGCAGCGGGGGCAGCCACACCCTT TCTCATCCACCTCCTTCGGGGGCTCCAGTTCTGGCAGCTCGGGGTTCAAGTCGGTGAT GTCATCCACCGAGATGATCAATGGCCACAAGGTCAACCAAGCGCATCGTGGAGAAC GGGCAGGAGCGCGTGGAGGTGGAGGAAGACGGGCAGCTCAAGTCGGTGACTGTGAACG GCAAGGAGCAGCTCAAATGGATGGACAGCAAGTAGGCGCTGGCCACCCGGCCCTGCCT TCCCACCACCAACCGTGATGGGGCAGCAAACAGTGGGGCCGAGACATAGCCTG ATGGTTAATAAATGT		
	ORF Start: ATG at 91	ORF Stop: TAG at 787	
	SEQ ID NO: 184	232 aa	MW at 25686.1kD
NOV63a, CG98131-01 Protein Sequence	MANYEVLGVQASAPEDIKKAYRKLALRWHPDKNPDNKEEAEEKFKLVSEAYEVLSD SKKRSLYDRAGCDWRAGGGASTPYHSPFDGTYTFRNPEDIFREFFGGLDPPSFEFWD SPFNSDRGGRGHGLRGAFAAGFPAFMEAFSSFNMLGCSGGSHHTFSSTSPFGSSS GSSGFKSVMSSTEMINGHKVTKRIVENGQERVEVEEDGQLKSVTVNGKEQLKWMDSK		

Further analysis of the NOV63a protein yielded the following properties shown in Table 63B.

Table 63B. Protein Sequence Properties NOV63a	
PSort analysis:	0.6400 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

5 A search of the NOV63a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 63C.

Table 63C. Geneseq Results for NOV63a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV63a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU17599	Novel signal transduction pathway protein, Seq ID 1164 - Homo sapiens, 267 aa. [WO200154733-A1, 02-AUG-2001]	1..232 36..267	231/232 (99%) 231/232 (99%)	e-136
AAM95431	Human reproductive system related antigen SEQ ID NO: 4089 - Homo sapiens, 267 aa. [WO200155320-A2, 02-AUG-2001]	1..232 36..267	231/232 (99%) 231/232 (99%)	e-136
AAW94066	Human DnaJ-like protein, HSPJ2 - Homo sapiens, 330 aa. [WO9855509-A2, 10-DEC-1998]	1..222 1..231	147/233 (63%) 179/233 (76%)	4e-78
AAV74126	Human prostate tumor EST fragment derived protein #313 - Homo sapiens, 317 aa. [DE19820190-A1, 04-NOV-1999]	1..222 2..232	147/233 (63%) 179/233 (76%)	4e-78
AAM40013	Human polypeptide SEQ ID NO 3158 - Homo sapiens, 309 aa. [WO200153312-A1, 26-JUL-2001]	1..221 1..225	106/232 (45%) 149/232 (63%)	6e-48

In a BLAST search of public sequence databases, the NOV63a protein was found to have homology to the proteins shown in the BLASTP data in Table 63D.

Table 63D. Public BLASTP Results for NOV63a				
Protein Accession Number	Protein/Organism/Length	NOV63a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9QY17	DnaJ homolog subfamily B member 8 (mDJ6) - Mus musculus (Mouse), 227 aa.	1..230 1..226	190/230 (82%) 203/230 (87%)	e-107
O75190	DnaJ homolog subfamily B member 6 (Heat shock protein J2) (HSJ-2) (MSJ-1) (HHDJ1) (MRJ) - Homo sapiens (Human), 326 aa.	1..222 1..231	147/233 (63%) 179/233 (76%)	9e-78
O54946	DnaJ homolog subfamily B member 6 (Heat shock protein J2) (HSJ-2) (MRJ) (mDj4) - Mus musculus (Mouse), 242 aa.	1..232 1..242	151/246 (61%) 185/246 (74%)	3e-77
Q8WN90	MOLECULAR CHAPERONE MRJ - Bos taurus (Bovine), 242 aa.	1..232 1..242	148/246 (60%) 180/246 (73%)	4e-77
O35723	DnaJ homolog subfamily B member 3 (DnaJ protein homolog 3) (Heat shock J3 protein) (HSJ-3) (MSJ-1) - Mus musculus (Mouse), 242 aa.	1..226 1..236	133/238 (55%) 169/238 (70%)	2e-65

PFam analysis predicts that the NOV63a protein contains the domains shown in the Table 63E.

Table 63E. Domain Analysis of NOV63a			
Pfam Domain	NOV63a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DnaJ	3..69	45/79 (57%) 62/79 (78%)	4.7e-40

Example 64.

- 5 The NOV64 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 64A.

Table 64A. NOV64 Sequence Analysis

Table 64A. NOV64 Sequence Analysis			
	SEQ ID NO: 185	2711 bp	
NOV64a, CG98164-01 DNA Sequence	CCTAGGCAACCCCGCTAAACAAGATGGCGACCTCCGAGAGGGCTCTCTGAGGACCAG AGCTGCCTCTCTCTGAGAGGCTTGGGCAGATCCCGAACTGGAGCCCGATCGTTACAG TTTCGCGCAGAAAAAGAGCGTCAGCCTTGCTGGTCTTTCCCATGGGCGAGAAGACG AAGGCAGCTCTAACATAGCCTCCTCTACCTGCTCCAGCAGCTCATGCACCCGCTATCA GGAGCTGGACTCGGACGAGATGAGGACCACGGCGAGGGCGAGGCGGGATCCGAGGAG TCCTCAGAGTCCGAAATGCTGAATTTGGAGGAGGAATTTGATGGGGTCTGAGAGAGG AGGCTGTGGCCAAAGCACTCCATCACTTGGGGCGCTCAGGCTCTGGGACTGAGCAAGT CTACCTCAATCTAATTTATCAGGTTGTAATTTAATTGATGTTAGCATTCTCTGTGGA TATGTTTCATCTACAGAAGTTGGATCTTTAGCGAATAAAATTTGAAGATTTATCTTGTG TGAGTTGTATGCCTTATCTCTAGAACTTAATGCTTCTCAAATAAATTTGACTACGTT CTTCAATTTCAAGCCACCAAAACCTCAAGAAGGCGGATTTTCCCAACAACCAAT TCTGAAATTTGTGATTTGTGAGCTATCATGCTCTCACTAACTAATTTGGATGGCA ATGAGATAGAAGAAATCAGTGGACTAGAGATGTGCAACAACCTAATTCACCTTAGTT GGCCAAACAATAAGATCAGGACAATTAATGGCTTAAACAAGTTACCAATCAAAATACTT TGTCTGAGCAATAACAGATTGAGATGATCAGAGTTTGGAGGATCTGAAAGCCCTGC AGAACCTGGATCTGTCCCAATCAGATAAGCAGCCTCCAGGGCTTAGAGAATCATGA CCTCCTGGAAGTGATCAACCTGGAGGATAATAAGATTGCTGAGCTGAGAGAAATAGAA TACATTAATAAATTTACCCATCTTCGAGTTCTCAATCTTCTAGAAAATCCAATTCAGG AAAAGTCTGAATATTGGTTCTTCGTAATTTTATGCTTCTGCGATTAAACAGAATTAGA TCAGAAGAAGATTAAAGTGGAAAGAAAGGTTTTCAGCAGTGAATAAATATGATCCTCC CCTGAAGTGGTTGCAGCTCAAGACCACCTGACCCATGTTGTCAACAGCGTGATGCAGC CGCAGAGGATCTTTGACAGCACTCTCCAGCCTGGATGCCCTTATCCCATGCTGAT ACTAGCTGGTCTGAAGCTTGTGGGAAACGAGAGCTTGCCCATCGCCTCTGCAGACAG TTTAGCACTTACTTCAGATATGGGGCTGTATACCACAAGACCACCTTACTTTGGAG AAGGGGATCGAGTTGATTATCATTTTATCTCTCAAGACGTTTTTGTAGAAATGGTGAA CATGGGAAATTCATTCTAACATTTAGTTATGGTAATCACAAGTATGGATTAAATAGG GACACCGTAGAAGGTATCGCAAGAGATGGTTTGGCAAGCTGTATTTCATATGGAATAG AAGGTGTAAGAAGTTTGAATATTCTCTATTTTGGAGCCTCGTTATATCTGGTGGTGCC CATGAACAGGAAAAATATGAGGGATATTTGCGGAGAAAAGGATTATTCACTCGTGCA GAAATTTGAATTTGCTGTCTCCAGAGTGGACCTTTATATTAAATTAATCAGAATTTTC CGGGATATTTTGTAGAGGTAATCAATGCAGATGATCTGGATGTTGCCTACCAAAACT GAGTCAGCTCATTAGAGAATACCTTGGATTGACTGAGGAACCTGCCAAGAGTTTGGCT ACAACCTGCAGATGTTAAGACATCCCACTGAAACCTGAAGCGCATCTCACAAGTATA TTTCTTCGAATATGGGTGATTTCCTGCATTCTACAGACAGAACTACCTCATTAAATTT TTGGGGCAAACCTTTAGCCAAAAAACACCGAGCGGAAAGAGATTCTATACACAGACAG CATGAGGCGAGCCCGGCAAGCTCTAATGGGAAGGATACGCCCTGATCAGACACTCTAT TTCAAAGGGGTCCAGTACCAGCACCTCTCACCAGTGGTCTACACTATTATACAACTTT AGAAGAACTCTGGAAGATTTTGTATCTTTGTGAAGACTATTTTAACTCCATTGGA CCATATCTGAAAAGAGTGGGAAGGATTCCTTGGTTTCCATGAAATGTTCAITGTTTC GGTTCTGTCCGTGGTCAAAGAATTGCCTTTCCAGCCTCCGAGGGGAGCATTTCTTC ACACCTAGGATCAGGAGCCAGTGACAGTGAGACCGAAGAGACCGGAAGCACTACCT ATACAATCATTTTCACATGAAAAAGAGTCTCACCACACAGACAACTCGGTCCAG TCATCAGTCGCCCAGGTTCCAACGTCAAACCCACCTCCCTCCAATCCCTCAGGCGCG CAGGTAGACTAGCACTTGATGTCTGATCCTAACATGGAACCTGCTCTGCTGATGTC GAATTCCTTGCCCTTACCTGGCCATGGGTCCAGCTGTTTCTCACTCAACCCATTACCA CGGAAGAATGTTCTACCTGCCCTTAATCTATCAGCCAGTTTCTCTGTGATTCTTTGG CTGGTGTCTTTAGTTTTTTAATTAATAAATTTGTTTCTTAAAA		
	ORF Start: ATG at 24	ORF Stop: TAG at 2499	
	SEQ ID NO: 186	825 aa	MW at 93626.3kD
NOV64a, CG98164-01 Protein Sequence	MATSERALLRTRAASLLRGLGRSRTGARSLOFRAEKERQPCWSFFPMGQKTKGSSNIA SYLLQQLMHRVQELSDGDEHDHGEAGESESESESEMLNLEEBFDGVLREEAVAKALH HLGRSGSGTEQVYLLNLTLSCNLDIVSILCGYVHLQKLDLSANKIEDLSVCSCMPYLL ELNASQNNLTFFNFKPPKNLKKADFHNQISEICDLSAYHALTKLILDGNEIEEISG LEM CNNLIHLSLANNKITTINGLNKLPKILCLSNQIEMTGLEDLKALQNLDSLHN QISSLQGLENHDLLEVINLEDNKIAELREIEYIKNLPILRLNLENPIQEKSEYWF VIFMLRLRLTELDQKKIKVEEKVSAVNKYDPPPEVVAQDHLTHVNVSMQPORIFDST LPSLDAPYPMILLAGPEACGKRELAHRLCRQFSTYFRYGACHTRPYFYGEGDRVDYH FISQDVFDENVNMKFIILTPSYGNHKGYNLNRDTEGIARDGLASCIHMEIEGVRSLKY SYFEPYRILVPMNKEKGYGLRRKGLFSRAIEFAVSRVDLYIKINQNFPGYFDEVI NADDLDVAYQKLSQIREYLGLETPAKSLATTADVKTSHLKPEAHPKTYISSNMGDF LHSTDRNYLIKFWAKLSAKKTPAERDSIHRQHEAARQALMGRIRPDHTLLFQRPVPA PLTSLGHYYTTLEELWKSFDLCEDYFKPPFGPYPEKSGKDSLVMKCSLFRFCPSKE LPFQPPGSGISSHLGSGASDSETEETRKALPIQSFSEKESHQHRQHSVPVISRPGSN VKPTLPPIPOGRR		

	SEQ ID NO: 187	2550 bp	
NOV64b, CG98164-02 DNA Sequence	AAGATGGCGACCTCCGAGAGGGCTCTCCTGAGGACCAGAGCTGCCTCTCTCTGAGAG GCTTGGGCAGATCCCGAATCGGACCGCATCGTTACAGTTTCGCGCAGAAAAAGAGCG TCAGCCTTGCTGGTCTTTTCCCATGGGGCAGAAGACGAAAGGCAGCTCTAACATAGCC TCCTCTACCTGCTCCAGCAGCTCATGCACCGCTATCAGGAGCTGGACTCGGACGGAG ATGAGGACCAGGGCGAGGGCGAGGCGGGATCCGAGGAGTCTCAGAGTCCGAAATGCT GAATTTGGAGGAGGAATTTGATGGGGTCTGAGAGAGGAGGCTGTGGCCAAAGCACTC CATCACTTGGGGCGCTCAGGCTCTGGGACTGAGCAAGTCTACCTCAATCTAATCTTAT CAGGTTGTAATTTAATTGATGTTAGCATTCTCTGTGGATATGTTCTATCTCAGAAGTT GGATCTTTCAGCGAATAAAATTGAAGATTATCTTGTGTGAGTTGTATGCCTTATCTC CTAGAACTTAATGCTTCTCAAATAAATTGACTACGTTCTTCAATTTCAAGCCACCCA AAAACCTCAAGAAGGGCGGATTTTTCCCAACAACCAATTTCTGAAATTTGTGATTTGTC AGCGTATCATGCTCTCACTAACTAATTTGGATGGCAATGAGATAGAAGAAATCAGT GGACTAGAGATGTGCAACAACCTAATTCACCTTAGTTTGGCCAAACAATAAGATCACGA CAATTAATGGCTTAACAAGTTACCAATCAAAATACCTTTGTCTGAGCAATAACAGAT TGAGATGATCACAGGTTTGGAGGATCTGAAAGCCCTGCAGAACCTGGATCTGTCCAC AATCAGATAAGCAGCCTCCAAGGCTTAGAGAATCATGACCTCTGGAAGTCAATCAACC TGGAGGATAATAAGATTGCTGAGCTGAGAGAAATAGAATACATTAAAAATTTACCCAT CCTTCGAGTTCTCAATCTTCTAGAAAATCCAAATTCAGGAAAAGTCTGAATATTGGTTC TTCGTAATTTTTATGCTTCTGCGATTAACAGAATTAGATCAGAAGAAGATTAAAGTGG AAGAAAAGGTTTCAGCAGTGAATAAATATGATCTCCCCCTGAAGTGGTTCAGCTCA AGACCACCTGACCCATGTTGTCAACAGCGTGATGCAGCCGAGAGGATCTTTGACAGC ACTCTTCCAGCCTGGATGCTCCTTATCCCATGCTGATACTAGCTGGTCTGTAAGCTT GTGGGAAACGAGAGCTTGCCCATGCGCTCTGCAGACAGTTTAGCACTTACTTCAGATA TGGGGCCTGTCTACACCAAGACCACCTTACTTTGGAGAAGGGGATCGAGTTGATTAT CATTTTATCTCTCAAGACGTTTTTGTGAAATGGTGAACATGGGGAAATTCATTCTAA CATTTAGTTATGGTAATCACAAATATGGATTAATAGGGACCCGTAGAAGTATCGC AAGAGATGGTTTGGCAAGCTGTATTATATGGAATAAGAAGTGTAGAAGTTTGAAA TATTCCTATTTGAGCCTCGTTATATCTCGTGGTGCCCATGAACAAGGAAAAATATG AGGGATATTTGCGGAGAAAAGGATTATTCAGTCGTGCAGAAATGAATTTGCTGTCTC CAGAGTGGACCTTTATATTAATAAATTAATCAGAATTTTCCGGGATATTTTATGAGGTA ATCAATGCAGATGATCTGGATGTTGCCTACCAAAAACCTGAGTCAGCTCATTAGAGAA ACCTTGGATTGACTGAGGAACCTGCCAAGAGTTTGCTACAACCTGCAGATGTTAAGAC ATCCCCACTGAAACCTGAAGCGCATCTACAAAGTATATTTCTTCAATATGGGTGAT TTCTCGCAATCTACAGACATAAACTACCTATTAATTTTGGCCCAAACCTTTCAGCCA AAAAAACACCAGCGGAAAGAGATTCTATACAGACAGCAGAGGCAGCCCGGCAAGC TCTAATGGGAAGGATACGCCCTGATCACACACTCTATTTCAAAGGGGTCCAGTACCA GCACCTCTCACCAGTGGTCTACACTATTATACAACITTAGAAGAACTCTGGAAGAGTT TTGATCTTTGTGAAGACTATTTTAAACCTCCATTGGACCATATCCTGAAAAGAGTGG GAAGGATTCCTTGGTTTCCATGAAATGTTTATTGTTTCGGTTCTGTCCGTGGTCAAAA GAATTGCCTTTCCAGCCTCCGAGGGGAGCATTCTTTCACACCTAGGATCAGGAGCCA GTGACAGTGAGACCGAAGAGACCCGGAAGCACTACCTATACAATCATTTTCACTGA AAAAGAGTCTCACCACACAGACAACACTCGGTCCCAGTCATCAGTCGCCAGGTTCC AACGTCAAACCCACCTCCTCCAATCCTCGGGGCGCAGGTAGACTAGCACTTGAT GTCTGATCCTAACATGAAAACCTGCTCTGCTGATGTCGAATTCCTTGCCTTACCT		
	ORF Start: ATG at 4	ORF Stop: TAG at 2479	
	SEQ ID NO: 188	825 aa	MW at 93602.3kD
NOV64b, CG98164-02 Protein Sequence	MATSERALLRTRAASLLRGLGRSRTGARSLOFRAEKERQPCWSFPMGQKTKGSSNIAS SYLLQQLMHRVQELDSGDGDEQGEAGESESESEMLNLEEEFDGVLREEAVAKALH HLGRSGSGTEQVYVNLTLSGCNLIDVSIILCGYVHLQKLDLSANKIEDLSVCVSCMPYLL ELNASQNNLTTFNFKPPKNLKKADFSHNQISEICDLSAYHALTKLILDGNEIEEISG LEMENNLIHLSLANNKITTINGLNKLPKILCLSNQIEMITGLEDLKALQNLDSLHN QISSLGLENHDLLEVINLEDNKIAELREIEYIKNLPILRVNLLENPIQEKSEYWF VIFMLRLTELDQKKIKVEEKVSAVNKYDPPPEVVAQDHLTHVNSVMQPRIFDST LPSLDAPYPMILLAGPEACGKRELALHRLCRQFSTYFRYGACHTRPPYFGEGRVDYH FISQDVFDEMVMNGKFIITFSYGNHXYGLNRDTEGIARDGLASCIHMEIEGVRSLKY SYFEPRYILVVPNMKEKEYEGLRRKGLFSRAEIEFAVSRVDLYIKINQNPFGYFDEVI NADDLDVAYQKLSQLIREYLGLETPAKSLATTADVKTSHLKPEAHPTKYISSNMGDF LHSTDINYLIKFWAKLSAKKTPAERDSIHRQHEAARQALMGRIRPDHTLLFORGPVPA PLTSGLHYTTLEELWKSFDLCEDYFKPPFGPYPEKSGKDSVSMKSLFRPCPWSKE LPFPQPEGSISSHLGSGASDSETEETRKALPIQSFSEKESHQHRQHSVPVISRPGSN VKPTLPPIPRGR		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 64B.

Table 64B. Comparison of NOV64a against NOV64b.		
Protein Sequence	NOV64a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV64b	31..825	750/795 (94%)
	31..825	751/795 (94%)

Further analysis of the NOV64a protein yielded the following properties shown in Table 64C.

Table 64C. Protein Sequence Properties NOV64a	
PSort analysis:	0.9074 probability located in mitochondrial matrix space; 0.6000 probability located in mitochondrial inner membrane; 0.6000 probability located in mitochondrial intermembrane space; 0.6000 probability located in mitochondrial outer membrane
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV64a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded
 5 several homologous proteins shown in Table 64D.

Table 64D. Geneseq Results for NOV64a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV64 a Residue s/ Match Residue s	Identities/ Similarities for the Matched Region	Expect Value
ABG07962	Novel human diagnostic protein #7953 - Homo sapiens, 1175 aa. [WO200175067-A2, 11-OCT-2001]	235..614 1..389	347/389 (89%) 354/389 (90%)	0.0
ABG07962	Novel human diagnostic protein #7953 - Homo sapiens, 1175 aa. [WO200175067-A2, 11-OCT-2001]	235..614 1..389	347/389 (89%) 354/389 (90%)	0.0
AAM05508	Peptide #4190 encoded by probe for measuring breast gene expression - Homo sapiens, 61 aa. [WO200157270-A2, 09-AUG-2001]	31..91 1..61	60/61 (98%) 60/61 (98%)	2e-28
AAM30367	Peptide #4404 encoded by probe for measuring placental gene expression - Homo sapiens, 61 aa. [WO200157272-A2, 09-AUG-2001]	31..91 1..61	60/61 (98%) 60/61 (98%)	2e-28
AAM17861	Peptide #4295 encoded by probe for measuring cervical gene expression - Homo sapiens, 61 aa. [WO200157278-A2, 09-AUG-2001]	31..91 1..61	60/61 (98%) 60/61 (98%)	2e-28

In a BLAST search of public sequence databases, the NOV64a protein was found to have homology to the proteins shown in the BLASTP data in Table 64E.

Table 64E. Public BLASTP Results for NOV64a				
Protein Accession Number	Protein/Organism/Length	NOV64a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96M69	CDNA FLJ32786 FIS, CLONE TESTI2002256, WEAKLY SIMILAR TO GUANYLATE KINASE (EC 2.7.4.8) - Homo sapiens (Human), 825 aa.	1..825 1..825	824/825 (99%) 824/825 (99%)	0.0
Q9D5S7	4921528H16RIK PROTEIN - Mus musculus (Mouse), 820 aa.	1..825 1..820	608/827 (73%) 680/827 (81%)	0.0
BAB85845	AXONEMAL LEUCINE-RICH REPEAT PROTEIN - Ciona intestinalis, 325 aa.	92..377 8..305	97/300 (32%) 160/300 (53%)	6e-31
O35125	CHROMOSOME 6 BAC-284H12 (RESEARCH GENETICS MOUSE BAC LIBRARY) COMPLETE SEQUENCE - Mus musculus (Mouse), 340 aa.	71..372 9..312	96/306 (31%) 162/306 (52%)	2e-30
Q99620	LEUCINE RICH PROTEIN - Homo sapiens (Human), 312 aa.	73..318 16..261	81/249 (32%) 130/249 (51%)	6e-23

PFam analysis predicts that the NOV64a protein contains the domains shown in the Table 64F.

Table 64F. Domain Analysis of NOV64a			
Pfam Domain	NOV64a Match Region	Identities/ Similarities for the Matched Region	Expect Value
LRR	238..259	8/25 (32%) 17/25 (68%)	0.21
LRR	281..302	10/25 (40%) 16/25 (64%)	0.69
Guanylate_kin	450..554	28/108 (26%) 66/108 (61%)	1.7e-07

Example 65.

- 5 The NOV65 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 65A.

Table 65A. NOV65 Sequence Analysis			
	SEQ ID NO: 189	723 bp	
NOV65a, CG99588-01 DNA Sequence	ATGGTGTCCCCGGTCACTGTGGTGAAGAGTGAAGGACCCAACTGGTGCCGTTCTTCAGGCCACCTGCGTGTATTTGTGCTCTGGCTGCCCTCATCTAGCCCATCGTGGGTGACACCCTCATCAAGTGCCTGCCTATCTTCTGCCTCTGGCTCTTCTCTTGGCCCATGGCTGGGATTCTCTGCTGGCCCAACCCAGCGCCACCCGCATCTTTGTGGGGCTTGTCTTCTCTGCTGTAGGTGACGCCTTCCTCATCTGGCAGGACCAAGGATACTTCGTGCATGGTCTGCTGATGTTTGTCTGTGACCCACATGTTCTACGCCTCGGCCTTGGCATGCAGCCACTGCTCTTCGGACAGGTCTGGTGTATGGCAGCGCTGTGCGGCCTGTGCTATGCCCTCCTCTACCCATGCCTCTCAGGTGCCTTACCTACCTGGTGGGGTCTATGTGGCCCTTATCGGCTTCATGGGCTGGCGAGCTATGGCAGGGCTGCGGCTGGCCGGGGCAGACTGGCGCTGACAGAGCTGGCAGCTGGCAGTGGTGCACCTCTTCTTTATCATCTCAGACCTGACCATCCCTCAACAAATTCTGTTTCTCTGTGCGCCTACTCTCGGGCGCTTATCATGTCCACCTCTATGCGGCCCAGATGCTCGTGCCTTGTGAGCTGTGCGAAGCCGGGAGCCTGTGGAACTACTACAGACTGACCAAGGCCAACTGA		
	ORF Start: ATG at 1	ORF Stop: TGA at 721	
	SEQ ID NO: 190	240 aa	MW at 26370.1kD
NOV65a, CG99588-01 Protein Sequence	MVSPVTVVKSEGPKLVPFFKATCVYFVLWLPSSSPSWVSTLIKCLPIFCLWLFLLAHGLGFLLAHP SATRI FVGLVFS AVGDAFLIWQDQGYFVHGLLMFAVTHMFYASAFGMQPIALRTGLVMAALSGLCYALLYPCLSGAFTYLVGVYVALIGFMGWRAMAGRLLAGADWRWTELAAGSGALFFFIISDLTIALNKFCFPVPYSRALIMSTYYAAQMLVALSAVESREPVEHYRLTKAN		

Further analysis of the NOV65a protein yielded the following properties shown in Table 65B.

Table 65B. Protein Sequence Properties NOV65a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.4000 probability located in Golgi body; 0.3000 probability located in endoplasmic reticulum (membrane); 0.0300 probability located in mitochondrial inner membrane
SignalP analysis:	Cleavage site between residues 65 and 66

A search of the NOV65a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 65C.

Table 65C. Geneseq Results for NOV65a

Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV65a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABG18549	Novel human diagnostic protein #18540 - Homo sapiens, 350 aa. [WO200175067-A2, 11-OCT-2001]	1..240 111..350	240/240 (100%) 240/240 (100%)	e-139
ABG18549	Novel human diagnostic protein #18540 - Homo sapiens, 350 aa. [WO200175067-A2, 11-OCT-2001]	1..240 111..350	240/240 (100%) 240/240 (100%)	e-139
AAM93259	Human polypeptide, SEQ ID NO: 2709 - Homo sapiens, 240 aa. [EP1130094-A2, 05-SEP-2001]	1..240 1..240	239/240 (99%) 239/240 (99%)	e-138
ABB59182	Drosophila melanogaster polypeptide SEQ ID NO 4338 - Drosophila melanogaster, 254 aa. [WO200171042-A2, 27-SEP-2001]	9..226 9..240	84/238 (35%) 118/238 (49%)	2e-26
AAY11618	Human 5' EST secreted protein SEQ ID NO:270 - Homo sapiens, 100 aa. [WO9906439-A2, 11-FEB-1999]	15..87 25..93	34/76 (44%) 47/76 (61%)	6e-10

In a BLAST search of public sequence databases, the NOV65a protein was found to have homology to the proteins shown in the BLASTP data in Table 65D.

Table 65D. Public BLASTP Results for NOV65a

Protein Accession Number	Protein/Organism/Length	NOV65a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q96AJ0	SIMILAR TO RIKEN CDNA 1810054O13 GENE - Homo sapiens (Human), 224 aa (fragment).	17..240 1..224	220/224 (98%) 220/224 (98%)	e-126
Q9D8N3	1810054O13RIK PROTEIN - Mus musculus (Mouse), 241 aa.	1..240 1..241	214/241 (88%) 224/241 (92%)	e-122
Q9VAJ0	CG7582 PROTEIN - Drosophila melanogaster (Fruit fly), 254 aa.	9..226 9..240	84/238 (35%) 118/238 (49%)	5e-26
Q9KLP0	HYPOTHETICAL PROTEIN VCA0703 - Vibrio cholerae, 229 aa.	21..225 31..228	68/215 (31%) 105/215 (48%)	1e-12
Q8YDD4	HYPOTHETICAL PROTEIN BMEII0243 - Brucella melitensis, 184 aa.	66..223 15..176	53/168 (31%) 79/168 (46%)	3e-11

PFam analysis predicts that the NOV65a protein contains the domains shown in the Table 65E.

Table 65E. Domain Analysis of NOV65a			
Pfam Domain	NOV65a Match Region	Identities/ Similarities for the Matched Region	Expect Value

Example 66.

The NOV66 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 66A.

Table 66A. NOV66 Sequence Analysis		
	SEQ ID NO: 191	1773 bp
NOV66a, CG99618-01 DNA Sequence	GGAGGAACATGACATCATGGAAATGGTTTCACCCAAATATCACTGGTGTGGAGGCAGA AAACCTACTGTTGACAAGAGGAGTTGATGGCAGTTTTTCGGCAAGGCCTAGTAAAAGT AACCTGGAGACGTCACACTTTCTGTTAGAAGAAATGGAGCTGTCAACCCACACCAAGA TTCAGAACACTGGTGATTGCTATGACCTGTATGGAGGGGAGAAGTTGCCACTTTGGC TGAGTTGGTCCAGTATTACATGGAACATCATGGGCAATTAAAAGAGAAGAATGGAGAT GTTATTGAGCTTAAAAATCCTCTGAACCTGTGCAGATCCTACTTCTCAAAGGTGGTTTC ATGGACACCTCTCTGAAAAGAAGCAGAGAAATTGTTAACTGAAAAGGAAAGCATAG TAGCTTTCTGTACGAGAGAGCCAGAGCCACCCTGATTTTGTCTCTCCGTGTGCACC GGTGATGACAAAGGAGAGCAATGACGGCAAGTCTAAAGTGACTCATGTCATGATTTC ACTGTCAGGAACGAAATACGATGTTGGTGGAGGAGAACGGTTTGATGCTTTGACAGA TCTTAAGAAGAATCCTATGGTGGAAACACTGGGTACAGTACTACAACTCAAGCAGCCC CTTAACACGACTTGTACAAATGCTGCTGAAATAGAAAGCAGAGTTTCGAGAACTAAGCA AATTAGCTGAGACCACAGATAAAGTCAGACAAGGCTTTTGGGAAGAATTGAGACACT GCAACAACAGGAGTGCAAACTTCTCTACAGCCAAAAGAGGGTTCGAAGGCAAGAAAAC AAAACCAAAAATAGATATAAAAACATCCTGCCCTTTGATCATAGGGTTGTCTACACA ATGGTGATCCAGTGAGGCTGTTTCAGATTACATCAACGCAGATATCATCATGCCTGA ATTTGAACCAAGTGCACAATTTGTAGCCAAAAAAGCTACTTGGGAACCGCCTGATT ACTGGGTTATTTCTTAATGAGTTCTGGCACATGGTTCTCATCTTTTAAAGAGGGAAAA GTGTAACTCGGCAGCACCAAGTGGAGAGAGGAAAGAGTAAATGTGTCAAATACTGGCC CGATGAGTATGCTCTAAAGAATATGGCATCGTGCATGTTAGGAACGTCAAAGAAAGC ACCACTCATGACTATAAGCTAAGAAAACCTAAACTTTCAAAGATTGGACAAGGGAATA TGGAGAGAATGGTCTGGCAATACCACTTGTGGACCTGGCTGTACCACGGAGTGCCCTG CGACCCCAAGGCATGCTGCACCTTCTGGACGAAGTGCACCATAGCAGGAGAGCATC ATGGATGCAGGGCTGGTCTGGTGGCCTGCAGTGTGGCATTGGCCAGACAGGGACGT TCATTGTGATTGATATCTTATTGACATCATCAGAGAGAAAGGTGTTGACTGCGAAT TGACGTTCCAAAACCATCCAGATGGTGCAGTCTCAGAGGTGAGAGATGGTCCAGACA GAAGCACAGTACCGATTATCTATACGGTGGTCCAGCATCATATTGAAACACTACAGC GCAGGATTGAGGAAGAGCAGAAAAGCAAGAGGAAAGGGCACGAATATACAAATATTAA GTATTCTCTAGTGGACAGACGAGTGGGGATCAGAGCCCCCTCTGCCTTATACTCCA ATGCCACCCTGTGCAAAATGAGAGAAGAGAGTCCAGAGTCTATGAAAACGTGGGCC TGATGCAACAGCAGAAACATTAAAGATGAGAAA	

	ORF Start: ATG at 9	ORF Stop: TGA at 1767	
	SEQ ID NO: 192	586 aa	MW at 67260.0kD
NOV66a, CG99618-01 Protein Sequence	MTSWKWFHPNITGVEAENLLLTRGVDGSFSARPSKSNPGDVTLSVRRNGAVTHTKIQN TGDCYDLYGGEKFATLAELVQYYMEHHGQLKEKNGDVIELKNPLNCADPTSQRWFHGH LSGKEAEKLLTEKGKSSFLVRESQSHPDFVLSVCTGDDKGESNDGKSKVTHVMIHCQ ELKYDVGGGERFDALTDLKKNPVETLGTVLQLKQPLNTTCTNAAEIESRVRELSKLA ETTDKVRQGFWEFETLQQQECKLLYSQKEGRRQENKTKNRYKNILPFDHRVVLHNGD PSEAVSDYINADIIMPEFEPSAQFVAKKSYLGTSLITGLFLNEFWHMLIFLRGKSVT RQHQVERGKSKCVKYWPDEYALKEYGIVHVRNVKESTTHDYKLRKLKLSKIGQGNMER MVWQYHLWTWLYHGVPCDPKGMHLFLDEVHKKQESIMDAGLVVVPCSAGIGQTGTFIV IDILIDIIREKGVDCEIDVPKTIQMVRSQRSEMVQTEAQYRFIYTVVQHHIETLQRRRI EEEQSKRKGEHTNIKYSLVDQTSQDQSPLLPYTPMPPCAEMREESARVYENVGLMQ QQKHLR		

Further analysis of the NOV66a protein yielded the following properties shown in Table 66B.

Table 66B. Protein Sequence Properties NOV66a	
PSort analysis:	0.4500 probability located in cytoplasm; 0.3719 probability located in microbody (peroxisome); 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV66a protein against the Geneseq database, a proprietary
5 database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 66C.

Table 66C. Geneseq Results for NOV66a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV66a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB59215	SHP-2 protein - Unidentified, 593 aa. [US6156551-A, 05-DEC-2000]	1..586 1..593	498/593 (83%) 521/593 (86%)	0.0
AAY13476	Peptide Seq ID No: 8 of WO9923493 - Homo sapiens, 593 aa. [WO9923493-A1, 14-MAY-1999]	1..586 1..593	498/593 (83%) 521/593 (86%)	0.0
AAR99313	Human SH-PTP2 (protein tyrosine phosphatase-2, with SH2 domains) - Homo sapiens, 593 aa. [US5536636-A, 16-JUL-1996]	1..586 1..593	498/593 (83%) 521/593 (86%)	0.0
AAB59221	SHP-2 activated double compound mutant E76A/D425A protein - Unidentified, 593 aa. [US6156551-A, 05-DEC-2000]	1..586 1..593	497/593 (83%) 520/593 (86%)	0.0
AAB59214	SHP-2 mutant E76A protein - Unidentified, 593 aa. [US6156551-A, 05-DEC-2000]	1..586 1..593	497/593 (83%) 520/593 (86%)	0.0

In a BLAST search of public sequence databases, the NOV66a protein was found to have homology to the proteins shown in the BLASTP data in Table 66D.

Table 66D. Public BLASTP Results for NOV66a				
Protein Accession Number	Protein/Organism/Length	NOV66a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q06124	Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48) (Protein-tyrosine phosphatase 2C) (PTP-2C) (PTP-1D) (SH-PTP3) (SH-PTP2) (SHP-2) - Homo sapiens (Human), 593 aa.	1..586 1..593	498/593 (83%) 521/593 (86%)	0.0
Q64509	PROTEIN TYROSINE PHOSPHATASE (EC 3.1.3.48) - Mus musculus (Mouse), 597 aa.	1..586 1..597	497/597 (83%) 522/597 (87%)	0.0
P41499	Protein-tyrosine phosphatase, non-receptor type 11 (EC 3.1.3.48) (Protein-tyrosine phosphatase SYP) - Rattus norvegicus (Rat), 593 aa.	1..586 1..593	494/593 (83%) 520/593 (87%)	0.0
Q90687	PROTEIN-TYROSINE PHOSPHATASE N11 (EC 3.1.3.48) (PROTEIN TYROSINE PHOSPHATASE, NON-RECEPTOR TYPE 11) - Gallus gallus (Chicken), 593 aa.	1..586 1..593	491/593 (82%) 518/593 (86%)	0.0
A53593	protein-tyrosine-phosphatase (EC 3.1.3.48), nonreceptor type 11 - rat, 597 aa.	1..586 1..597	493/597 (82%) 519/597 (86%)	0.0

PFam analysis predicts that the NOV66a protein contains the domains shown in the Table 66E.

Table 66E. Domain Analysis of NOV66a			
Pfam Domain	NOV66a Match Region	Identities/ Similarities for the Matched Region	Expect Value
SH2	6..81	39/79 (49%) 66/79 (84%)	8e-38
SH2	112..192	33/84 (39%) 72/84 (86%)	3.6e-29
Y_phosphatase	268..513	98/298 (33%) 187/298 (63%)	1.7e-58

Example 67.

The NOV67 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 67A.

Table 67A. NOV67 Sequence Analysis			
	SEQ ID NO: 193	1311 bp	
NOV67a, CG99832-01 DNA Sequence	CCGGGATGCACTGTTCTGCTGTGGGTCTCATCATGGAGACCAACGGGTGGAGATT CCCGGCAGCGTCTTGGACGATTCTGCAGCCGATTATTTTGCATATTTCCAGCGAGG AAAGAGACAATGCAATCCGAGTGTGTTTTAGATTGAACTTGCCCATGGTTTACTT GGATTTCTACATGCAGAACACACCAGGATTACCTCAGTGTGGGATAAGAGACTTTGCT AAAGCTCTCTTCAGTCATTGTCCGTTTTTGTGCTGCCTCAAGGTGAAGATGTGGAAAAAC TTTTGGATGAATGGAAGGAATATAAAATGGGAGTACCAACATATGGTGCAATTATTCT TGATGAGACACTTGAAATGTACTACTAGTTTCAGGGGTACCTAGCAAAATCAGGCTGG GGATTTCCAAAAGGAAAAGTAAATAAAGAAGAAGCTCCTCATGATTGTGCTGCTAGAG AGGTATTTGAAGAACTGGTTTTGATATCAAAGACTATATTTGTAAGGATGATTACAT TGAATTCGAATCAATGACCAGCTTGCTCGTTTGTACATCATTCAGGAATCCAAA GACACAAAATTTAAACCAAACTAGAAGAGAAATTCGGAACATTGAGTGGTTCTCTA TTGAGAAATTGCGCTTGTCATAGAAATGATATGACCCCAATCCAAACTTGGTTTGGC ACCTAACAAATTTTTATGGCCATTCCCTTTATCAGGCCATTAAAGGACTGGCTTTCT CGAAGATTTGGCGATTCTCAGACAGTGACAATGGATTTTCTCAACTGGTAGCAGC CGGCTAAACCCACTGTGGAAAATTGAGGAGTCGAACCAATTCGCCACAGTCAGCA GTTATTTCTGACGGTTCTCCTGGTGACCAGTGGGTAAAGCACAGGCAACCACTGCAG CAAAAGCCATATAATAATCATTCTGAAATGTCTGACCTTTTAAAGGAAAGAATCAA GTATGAGGGGAAATGGCAGAAAACAGTATCAAGATTCACCTAATCAAAAGAAAAGAA AAATGGGCTTCAGCCAGCAAAGCAGCAGAAATCTTTGATGAAGTGTGAAAAGAACTT CATCCACGGAACCTTCAGGATAATTTTGAACAGATGCTGTATATGACTTGCCTAGCT CCAGTGAAGACCACTTGCTAGAACATGCTGAGGGACAGCCCGTGGCATGTAATGGACA TTGCAAGTTCCTCTTTTCATCCAGAGCCTTTTGAGTTTCAAGTTTGACCATAATGCT ATAATGAAAATCTTGGACCTTTGATAGCAGCACAT		
	ORF Start: ATG at 35	ORF Stop: TGA at 1298	
	SEQ ID NO: 194	421 aa	MW at 48626.8kD
NOV67a, CG99832-01 Protein Sequence	METKRVEIPGSVLDDFCRSFILHIPSEERDNAIRVCFQIELAHWFYLDFFYMONTFGLP QCGIRDFAKALFSHPFLLPQGEDVEKVLDEWKEYKMGVPTYGAILDETLENVLLVQ GYLAKSGWGFPGKVNKEEAPHDCAAREVFEETGFDIKDYICKDDYIELRINDQLARL YIIPGIPKDTKFNPKTRREIRNIEWFSIEKLPCRNDMTPKSKLGLAPNKFMAIPFI RPLRDWLSRRFGDSSDSDNGFSSTGSTPAKPTVEKLRSRTKFRHSQQLFPDGPSPGDQ VKHRQPLQKFPYNNHSEMSDLLKGNQSMRGNRKYQDQSPNQKIRTNGLQPAKQONS LMKCEKKLHPRKLQDNFETDAVYDLPSSEDLLEHAEGQPVACNGHCKFPFSSRAFL SFKFDHNAIMKILD		

Further analysis of the NOV67a protein yielded the following properties shown in

5 Table 67B.

Table 67B. Protein Sequence Properties NOV67a	
PSort analysis:	0.4741 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV67a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 67C.

Table 67C. Geneseq Results for NOV67a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV67a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
ABB62134	Drosophila melanogaster polypeptide SEQ ID NO 13194 - Drosophila melanogaster, 791 aa. [WO200171042-A2, 27-SEP-2001]	4..261 205..474	131/270 (48%) 171/270 (62%)	4e-72
AAM90822	Human immune/haematopoietic antigen SEQ ID NO:18415 - Homo sapiens, 95 aa. [WO200157182-A2, 09-AUG-2001]	253..348 1..92	91/96 (94%) 92/96 (95%)	8e-46
AAW87487	S. cerevisiae TIH1 polypeptide - Saccharomyces cerevisiae, 595 aa. [US5846764-A, 08-DEC-1998]	12..239 18..235	82/228 (35%) 125/228 (53%)	2e-36
AAW54167	S.cerevisiae TIH1 protein - Saccharomyces cerevisiae, 595 aa. [WO9813502-A2, 02-APR-1998]	12..239 18..235	82/228 (35%) 125/228 (53%)	2e-36
AAR81314	Yeast TIH1 - Saccharomyces cerevisiae, 595 aa. [WO9519988-A1, 27-JUL-1995]	12..239 18..235	82/228 (35%) 125/228 (53%)	2e-36

- 5 In a BLAST search of public sequence databases, the NOV67a protein was found to have homology to the proteins shown in the BLASTP data in Table 67D.

Table 67D. Public BLASTP Results for NOV67a				
Protein Accession Number	Protein/Organism/Length	NOV67a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9CYC6	5730537H01RIK PROTEIN - Mus musculus (Mouse), 422 aa.	1..421 1..422	387/423 (91%) 402/423 (94%)	0.0
Q9VUU4	CG6169 PROTEIN - Drosophila melanogaster (Fruit fly), 791 aa.	4..261 205..474	131/270 (48%) 171/270 (62%)	1e-71
O13828	HYPOTHETICAL 83.2 KDA PROTEIN C19A8.12 IN CHROMOSOME I - Schizosaccharomyces pombe (Fission yeast), 741 aa.	12..242 11..236	109/231 (47%) 153/231 (66%)	1e-59
Q948G1	PUTATIVE ABC TRANSPORTER - Oryza sativa (Rice), 1372 aa.	18..265 114..368	96/258 (37%) 148/258 (57%)	9e-46
E88886	protein F52G2.1a [imported] - Caenorhabditis elegans, 328 aa.	4..195 129..323	89/195 (45%) 123/195 (62%)	8e-44

PFam analysis predicts that the NOV67a protein contains the domains shown in the Table 67E.

Table 67E. Domain Analysis of NOV67a			
Pfam Domain	NOV67a Match Region	Identities/ Similarities for the Matched Region	Expect Value
NUDIX	96..221	25/137 (18%) 78/137 (57%)	4.7e-12

Example 68.

- 5 The NOV68 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 68A.

Table 68A. NOV68 Sequence Analysis

	SEQ ID NO: 195	5474 bp
NOV68a, CG99842-01 DNA Sequence	TGATCATAAACAATACTTTTGGAGCTGGATTAATATTAAGTCACCTGAGAGATATAGA AAATACCATGTTCTGGTCAACTTTGATTATTTAGCCGGGTTTCATCTGGTTTGACACA TTGAGACCGTAATACTTCAGGTTCTAATTGTGTCCAGAAGGCTATGTGTTAACATTCA TAGCAAACCTGGTTACCTTTATTTCTGAAAGTTAAATTTTGAAGAAAGGGATACCAA GTGAGAAAGCAACCTTCAGTTAAACAAAGGGGTACATCAAGTTTGATACTCTTGATT AAAAATCAGGTGACAATTGGAACATGGAGGAGAAATTCATGATTATAGAAACCATGAA GCCAATCTTTTACAAACAAAGCTGCTTGTCTTTTCTTAATCATGTAGTGGCTGCTA TCAGGCCAGTTCTGGAACCTCAGTGAAGTCCCTCTGCTCTTCTGAATCAGCTCTTGATT GGTGGTGACTTCTCGGAAGTCTTATCCCTACCTCCACTCTGCTCCAGGGTGGCA CGACTCAAACTGCCCACGGCCATTAGCTCCATTCCATGAAGTTCCTATCCAGTAC CAGGCACGTTCTTCTCGCTCTGCTGGCTCGCGTGTCTTTTTCCTGTGTCCCGC CGGGTCAGGCTCTCCCGCTCTCTCAGACTTTCCCGTCCCGGTGCTCTCCGCGCGCTC TCTCTCAACTCTCTCCCGCTCTCAGCCACTTCTGCTCCAGTATGCTCCCTGGCGCG GCGCGCTGGCGCTCTGCGCACCTCTCTCCCGTCCGCTCTTCTTAGGAAAGCTGAG CCTCATAGCTTCCGGGAGAAGGTTTCCGGAAGAACTCCAGTCTGTGCAGTATGTA AGGTGACCATCGATGGGACAGGCGTTTCTGTGAGAGTCTGCAAGGTGGCGACGACAG AAAATGTGAAGCAAAGGTGACTTCAGCCTGTGAGCCTTGCTCCCGTGGAGTTGCGG CGAAACAGCGCCCTCAGTCAGGCGCATAGAGCACCTGGGATCCACAAATCTCTGAACC ACTCAAAGCAGCGCAGCACTCTGCCAGGAGCTTCAAGCCTGGACCCGCTCATGGAGCG GCGCTGGGACTTAGACCTCACCTACGTGACGGAGCGCATCTGGCGCGCGCTTCCCG GCGCGCCCGATGAAACAGCGGACCGGGGCCACCTGCGCGAGCTGGCCCATGTGCTGC AATCCAAGCACCGGGACAAGTACCTGCTCTTCAACCTTTAGAGAAAAGGCATGACCT GACCCGCTTAAACCCCAAGGTTCAAGACTTCGGCTGGCCTGAGCTGCATGCTCCACCC CTGGACAAGCTGTGCTCCATCTGCAAAGCCATGGAGACATGGCTCAGTGTGACCCAC AGCAGTGGTCTGTACTATATGCAAGGAAACAAGGGCAAGCTTGGGGTCACTGTTTCT TGCCTACATGCATACAGCAAGATCTCTGACGGGGCGGACAGGCACTGGCCACTCTT ACCATGCGGAAATCTGCGAGGACAAGGTGGCCACAGAATGACAGCCCTCCAGCGCT GATATATCAGCTACTTCAGTGGGCTGCTATCTGGCTCCATCAGAAACAGCAGCCCT TCTCTTCTGCATATGTGCTCATCCCCATGCTGCCAGCCTTTGAACCTGGCACAGGC TTCCAGCCCTTCTTAAATCTACAGTCCATGCAGCTTGTCTACACATCTGGAGTCT ATCATTGTCAGGCTTGGTCCCCAGCAGCTTGCATCAGCCTGGAGCCAGCCCTCTCT CCTCAAAGGCGATGTCTAGGTAACATGTTATCACAAGGGTGGCCGGGACAGACCGG ACCCTCGTGTTCGAGTCCAGTTCACACCTGCACCATCCACGGACACAGCTCACTT TCCCCAAGGACAGCTTGACGAGGCTGGACTGATGAGAGGTTCCCTTCCAAAGCCTC CGTGGAGTTTGTCTTCTCTCCAGCCCGAGAAGATCAAAGGCAAGCTCCACGGAAC GACCCCTCGGTCTCTGTGACTACAACCACTGAGCCAGCCGTGCGCTGGGACTCCT ATGAGAACTTCAACCAGCACACGAGGACAGTGTGGATGGCTCCTGACCCACACCCG GGTCCCCCTGGATGGCAGTCTTATGCCAGGTGACGCGCTCCCCGGCAGACCCCG CCGGCACCTCTCCAGAGCTCCACACCCCTATGCTCTCTGTGACAGCAGTCCAG GCCATTCTCCAGCTGACCACAGAGCCGCTGCTGAGTCCCTGGCGCGCGCCCCC TACAGTGTCTGAACGGCAGGAGCTGGATCGCTCTAGGAGGCTGGCGAGTGGCCAGT GGGGCGCGGAGCTGGCGCGAGACGGCCATCTAGATGACGAAGAGCAGCCCACTG TGGGCGGAGGCCCACTCGGAGTGTATCCAGGCCATAGGCTGGCTCAGCCGCA CTGCTCTGCGCGCAGGCTACCGGGAGCCTGCGGGGTCCCAATGGGGGCTACTAC CGGCCAGAGGGAACCTGGAGAGGAGGCGACTGGCTACGGGGCTATGAGGGATGCC CCCAGGGCTACGCCGAGGCTCGATGGAGAAGAGGCGCTCTGCCGATCGTGTGAGA GGGCTATACCCCTACCCACCTGAGATGGGGAACAGCCACTGGGGACTTTGGCTAC CGCGCCCCAGGCTACCGGAGGTGGTCATCTGGAGGACCTGGGCTGCTGCCCTAT ACCATGCCAGCCTGCGAGGAGAAGCTGGCGCTGCCTACAGCAGCCTGTATGGACT GCGGCTGGAGAGGAGCTGGAGAAGGGTGGGCAAGTGAGGCTGGCAAGCCTCTCCTG CACCCAGTGGCGCTGGGCAACCGCTGCTCTGCTCTTGCCTGCTGTGGGCATCACC ATGCCCGATGCTGACTACAGTGCCTGAAGCCACCAAGGCAGGCGAGGAAGGCA CGAGGGCTGCTCTACACATGTGCCCCGAAGGCAGGTATGGCATCCAGGTACCTT GCCCTGGTGACATACAGTATGGAGGAGCAGTTCAGTACTTGCAGCATATGGCC GTGTGCTCATAGCTGTGGCTCTCCAGGAGAGGCGAGAGGTATCCAGCCTGGTGC CCACTCCCCACGGGCTGGCTCCATTTCCCGGGCAGCCGCGCTATCCAAATCTAGG AAGCTGAGCTACGAGATCCCTACGGAGGAGGAGGGGACAGGTACCCATTGCTGGGC ACCTGGCCTCAGCAGGACCTTTGGCATCTGCAGAGTCTGTGGAGCCGGTGTCTGGAG GGAGGGCCCCAGTGGGCACAGCACTGCTCGGTCTCCCGAGATGCCCATGCACT GCTTCGTGAGAGTTGTGTGCTCCCTCCAGCCCTGCAACAGCAGTCCAGTCCAGG GCAAGGAAGCACCCGGCGACAGGACACAGGTCCCCACCTCAGCCCTCAGAG ACTGAGTCTGGCGAGGCTTGGCCCCCTGTTCCAGGAGGCAACGGAAAGGCCCT GAGTGGCGTGGGAAGTGGGCTGAGCTCTGGCCCCTAGCCAGTCTCTCCGACT TCCCTCCAGCTCGCCAGTGAAGTGGCTCAGGAAAGGAGTCCAGGGGGCTCAGAG TGGCGCAGTCTCGGAGCCTGTGCCCCACCACTTCTGGCTCCGCCAGCCCCC TGGCAAGGCCCTCGAGGCCCCCCGACAGCCAGATGGGTCTCCCTCACTCTGTGC CTCCAGATGCCCTGGCTTGTGGCCAGCCAGAGCCGCTCAGAGTCACTACACC	

	TGCTTTCCCTGGCTGCCTCCTATGACACCAATGGCCTTAGCCAGCCCCACTTCTCT GAGAAACGCCACCTGCCCGGGCCGGGGCAACAGCCAGGACCCCTGGGGCCAGAGCAGG CATCATCGCCAGCCAGAGGCATCAGTCACCATGTACCTTCGCACCTCTGCTCTCAGA TAATGTCCCCCAACCCAGAGCCTCCTACACAGAGAGCCAAAGCAATGTCAAGTTT GTCCAGGATACATCAAGTTCTGGTACAAGCCACACCTGTCCCGTGACCAAGCCATGT CCCTGTGAAGGACAGGACCCCTGGGGCCTTCTGTATCAGGGACAGTCATTATTCCA AGGAGCTTATGGGCTGGCCCTCAAGGTGGCCACACCGCCACCCAGTGCCAGCCCTGG AAAGGGGACCCCGTGGAACAGCTGGTCCGCCATTTCTCATCGAGACTGGGGCCAAAG GGGTGAAGATCAAGGGCTGCCCCAGTGAGCCCTACTTTGGCAGCCTGTCCGCCTTGGT CTCCAGCACTCCATCTCCCCATCTCCCTGCCCTGTGCTGCGCATTCACAGCAAA GATCCTCTGGAAGAGACCCAGAGGCTCCAGTGCCCAACCAATGAGCACAGCGGCAG ACCTCCTGCGTCAGGGTGCTGCCTGCAGCGTGCTCTACTTGACCTCAGTGGAGACAGA GTCACTGACGGGCCCCAAGCTGTGGCCCGGGCCAGCTGTGACGCTTGAGCTGTAGC CCCCGCCGACACCAAGCTGTTGTCACTTCAAGGTGTGACGCCAGGCTTACACTGA CGGACAACCAAGGAAGCTCTTCTTTCGCGCCATTATCCAGTGAACAGCATCACCTT CTCCAGCACTGACCTCAAGACCGGAGATGGACCAACCCAGAGCGGACCACTCCAG ATCTTTGGTTTCGTGGCCAAAGACCGGGAAGCCCTGGGAGAATGTGTGTACCTCT TTCAGAGCTTGACCCAGATCAGCCTGTGGCGCCATTGTACCTTTCATCACCAGTGT TCTACTGGGCCAGAGAAAATGAAGGAAGGCCACAAGCTCAGAGCCACATCAACTG CCCCCTCCAGCACCCACAGCCCTCACATCCCTGGCCTGAGCCAGGAGACCCAG GAGAAAGCACCTCCCTTAGGAATGAGGAGTGGGCATCAGGCCTGGGACACTGCTCTC CTTCCCGCCCCCAGCTGCTAAGTTAAGTGGACAGGCCACACAGATGACCTTGCAATG TGAGCAGATGGCAGAGATGGGTGTGTGAGGGGTGAGGAGGCATCAGCAGTTGAGCCCC GAAGGAGATCAGGCAGCCCCACCTGCAGGAGAAGCTCAGCCCTCCAGGGGATCAGCCC CTGCCAGTTCACCCAGCTGCAGGTGCCAGCACGGCAGGATGGGAGAGGGGTGGGGA GCGAGTCACTGCCTCCTCTGAGCAGAGATTAGAGTAGGATCACATGAATAGGGGAAA AAAGAGAGTCTATTTTGTCTAATAATAAAGAATTTCTATAAATTTAAAAA AAAAAAAAAAAAAAAAAAAA		
	ORF Start: ATG at 561 ORF Stop: TGA at 4950		
	SEQ ID NO: 196	1463 aa	MW at 158629.7kD
NOV68a, CG99842-01 Protein Sequence	MKFLSSTRHVSSASVWLACFLPVSRRLVRLSPSLRLSRPVLGRSLSPSLSPQLPV SSMSLARPGWSPAHPLSPRLFPRKAEPHSFREKVFRKKPVCAVCKVTIDGTGVSCRV CKVATHRKCEAKVTSACQALPPVELRRNTAPVRRIEHLGSTKSLNHSKQRSTLPRFS LDPLMERRWDLDTYVTERILAAFPARPDEQRHRGHLRELAHVLSKHRDKYLLFNL SEKRHDLTRLNPKVQDFGWPELHAPPLDKLCSICKAMETWLSADPQHVVVLYCKGNKG KLGIVISAYMHYSKISAGADQALATLTMRKFCEDKVATELQPSRRYISYFSGLLSGS IRMNSSPLFLHYVLIPMLPAFEPGTGFQPFLLKIYQSMQLVYTSVYHIAGPGPQQLCI SLEPALLKGDVMTVCYHKGGRGTDRTLVRVQFHTCTIHGPQLTFPKDQLEAWTDE RFPFQASVEVFSSSPEKIKGSTPRNDPSVSDYNTTEPAVRWDSYENFNQHHEDSVD GSLTHTRGPLDGSFYAQVQRPQTPPAPSPPEPPPPMLSVSSDSGHSSTLTTEPAAB SPGRPPPTAAERQELDRLLGGCGVASGGRGAGRETAILDDEEQPTVGGGPHLGVYPGH RPLSLRHCSRCQGYREPCGVPNGGYRPEGLTERRRLAYGGYEGSPQGYAEASMEKRR LCRSLSEGLYPYPPPEMGKPATGDFGYRAPGYREVILEDPGLPALYPCPACEKLALP TAALYGLRLEREAGEGWASEAGKPLLHPVRPGHPLPLLLPACGHHHAMPDYSCLKPP KAGEEGHEGCSYTMCEGRYGHGYPALVTYSYGGAVPSYCPAYGRVPHSCGSPGEGR GYPSPGAHSRAGSISPGSPYPQSRKLSYEIPTEEGDRYPLPGHLASAGPLASAES LEPVSWREGPSGHSTLPRSPRDAPCSASSELGSPSTPLHTSSSPVQKSTRQDTRSP TSAPTQRLSPGEALPPVSOAGTGKAPELPSGSGPEPLAPSPVSTFPSPSPDWQER SPGGHSDGASPRSPVPTTLPLRHAPWQGRGPPDSPDGSPLTPVPSQMPWLVASPEP PQSSPTPAPPLAASYDTNGLSQPPLPEKRHLPGPGQPGPWGPEQASSPARGISHHVT FAPLLSDNVPTPEPTQESQSNVQVQDTSKFWYKPHLSRDQAIALLKDKDPGAFLI RDSHSFQGAYGLALKVATPPPSAQPKGDPVEQLVRHFLIETGPKGVKIKGCPSEPYF GSLALVSQHSISPIPLPCLRIIPSKDPLEETPEAPVPTNMSTAADLLRQGAACSVLY LTSVETESLTGPQAVARASSAALSCSPRPTPAVVHFKVSAQGITLTDNQKRLFFRRHY PVNSITFSSTDQDPRRWTPDGTTSKIFGFVAKKPGSPWENVCHLFAELDDPDQFAGAI VTFITKVLLGQRK		

Further analysis of the NOV68a protein yielded the following properties shown in Table 68B.

Table 68B. Protein Sequence Properties NOV68a	
PSort analysis:	0.3700 probability located in outside; 0.1900 probability located in lysosome (lumen); 0.1080 probability located in nucleus; 0.1000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	Cleavage site between residues 28 and 29

A search of the NOV68a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 68C.

Table 68C. Geneseq Results for NOV68a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV68a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAB40638	Human ORFX ORF402 polypeptide sequence SEQ ID NO:804 - Homo sapiens, 1400 aa. [WO200058473-A2, 05-OCT-2000]	64..1463 1..1400	1400/1400 (100%) 1400/1400 (100%)	0.0
AAM40312	Human polypeptide SEQ ID NO 3457 - Homo sapiens, 1409 aa. [WO200153312-A1, 26-JUL-2001]	80..1463 26..1409	1383/1384 (99%) 1383/1384 (99%)	0.0
AAU17196	Novel signal transduction pathway protein, Seq ID 761 - Homo sapiens, 524 aa. [WO200154733-A1, 02-AUG-2001]	942..1463 3..524	522/522 (100%) 522/522 (100%)	0.0
AAU17573	Novel signal transduction pathway protein, Seq ID 1138 - Homo sapiens, 522 aa. [WO200154733-A1, 02-AUG-2001]	942..1463 1..522	521/522 (99%) 521/522 (99%)	0.0
AAB42301	Human ORFX ORF2065 polypeptide sequence SEQ ID NO:4130 - Homo sapiens, 523 aa. [WO200058473-A2, 05-OCT-2000]	149..582 48..492	232/445 (52%) 311/445 (69%)	e-131

- 5 In a BLAST search of public sequence databases, the NOV68a protein was found to have homology to the proteins shown in the BLASTP data in Table 68D.

Table 68D. Public BLASTP Results for NOV68a				
Protein Accession Number	Protein/Organism/Length	NOV68a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
Q9UPS7	KIAA1075 PROTEIN - Homo sapiens (Human), 1400 aa (fragment).	64..1463 1..1400	1399/1400 (99%) 1399/1400 (99%)	0.0
Q96P25	TENSIN2 - Homo sapiens (Human), 1285 aa.	179..1463 1..1285	1284/1285 (99%) 1284/1285 (99%)	0.0
Q9NT29	HYPOTHETICAL 68.7 KDA PROTEIN - Homo sapiens (Human), 649 aa (fragment).	813..1463 1..649	648/651 (99%) 648/651 (99%)	0.0
AAH25818	SIMILAR TO KIAA1075 PROTEIN - Mus musculus (Mouse), 655 aa (fragment).	808..1463 2..655	564/663 (85%) 580/663 (87%)	0.0
A54970	tensin, cardiac muscle - chicken, 1744 aa.	149..1140 31..1088	382/1111 (34%) 513/1111 (45%)	e-139

PFam analysis predicts that the NOV68a protein contains the domains shown in the Table 68E.

Table 68E. Domain Analysis of NOV68a			
Pfam Domain	NOV68a Match Region	Identities/ Similarities for the Matched Region	Expect Value
DAG_PE-bind	86..133	16/51 (31%) 33/51 (65%)	3e-05
PHD	99..136	9/51 (18%) 20/51 (39%)	0.22
SH2	1194..1286	28/97 (29%) 66/97 (68%)	5.4e-15

Example 69.

- 5 The NOV69 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 69A.

Table 69A. NOV69 Sequence Analysis			
	SEQ ID NO: 197	1502 bp	
NOV69a, CG99944-01 DNA Sequence	<div>GAATAAGGAGGAAGAATCATGAATGAAAAAGTCATAGAAATGCAGAATATTACCAAAG AATTTGGAACATTCAAGGCAAATGACAACATCAATTTACAAGTAAAGGCAGGAGAGAT TCATGCGTTGCTTGGAGAAAACGGTGTGGCAAATCTACATTGATGAACGTGCTTTCC GGATTATTAGAGCCGACATCAGGGAAAATTTTGATGCGTGGGAAAGAAGTACAGATCA CAAGCCCGACAAAAGCCAATCAATTAGGGATTGGGATGGTCCATCAGCACTTTATGCT TGTGATGCGCTTTACTGTAACAGAAAACATCGTGTGGGAAGCGAACCTAGTCGTGCA GGGATGCTTGACCATAAAAAAGCGGAAAAGAGATCCAAAAAGTTTCTGAACCAATATG GATTATCAGTCAACCCGGATGCTTATGTTCTGTGATATTTCAAGTGGGATGGAACAACG GGTAGAAATTTTAAAAACACTTTACCGAGGAGCAGATGTAAGTATTTTGTATGAGCCG ACAGCTGTATTGACCCCTCAGGAAATTGATGAATTAATCGTGATCATGAAGGAATTAG TCAAAGAAGGCAAGTCAATCATTGATTACGCATAAGTTAGATGAAATCAAAGCAGT AGCTGACCGTTGTACAGTTATCCGCGGTGGAAAAGGAATCGGTACAGTCAACGTTAAG GACGTTACCTCACAGCAATTAGCTGATATGATGGTCGGAAGACCGGTTTCATTCAAAA CGATGAAAAAGAAGCGAAGCCTCAAGAAGTCGTTTGTCTATTGAAAATCTAGTGGT AAAAGAAAATCGTGGATTAGAAGCCGTGAAAACCTGAACCTAGAGGTTCTGTCTGGC GAAGTACTTGGTATCGCTGGAATCGATGGAACCGGGCAGTCGGAGTTGATCCAAGCTT TGACTGGTTTTCGCAAGGCAGAAAGCGGACATATCAAGCTAAAAGGGGAAGACATCAC CAATAAAAAACCTCGAAAGATCACTGAACATGGTGTAGGACATGTGCCAGAAAGCCGT CATAAATACGGGTTGGTCTAGATATGACATTGTCTGAAAACATTGCCCTGCAAAACGT ATCATCAAAAACCTTACAGTAAAAACGGTATGCTGAATTATTCAAGTGATAAATGAACA TGCCAGAGAATTGATCGAAGAATATGATGTTTCGAACAACGAATGAACCTGTTCTGCA AAAGCTTTATCAGGCGGAAATCAGCAAAAAGCAATCATCGCTCGGATAGTCGACCGAG ATCCTGATCTGTTGATCGTTGCAAATCCAACCTCGTGGGCTGGATGTAGGTGCGATCGA ATTTATTCTAAACGCTCTGATCGAACAAGGGACAAATACAAAGCAGTGTATTGATT AGTTTCGAATTAGAAGAAATTTTAAATGTTTCGGATCGTATTGCTGTTATCCATGAAG GAGAAATCGTCGGGATCGTTGATCCGAAAGAAACATCTGAAAATTAATAGGG</div>		
	ORF Start: ATG at 19	ORF Stop: TAA at 1495	
	SEQ ID NO: 198	492 aa	MW at 54439.3kD
NOV69a, CG99944-01 Protein Sequence	<div>MNEKVIEMQNITKEFGTFKANDNINLQVKAGEIHALLGENGAGKSTLMNVLSGLLEPT SGKILMRGKEVQITSPTKANQLGIGMVHGFMLVDAFTVTENIVLGSEPSRAGMLDHK KARKEIQKVSQYGLSVNPDAYVRDISVGMEQORVEILKTLRYGADVLI FDEPTAVLTP QEIDELIVIMKELVKEGKSIILITHKLDEIKAVADRCTVIRRGKIGITVNVKDVTSQQ LADMMVGRAVSFKTMKKEAKPQEVVLSIENLVVKENRGLAEAVKNLNLVVRAGEVLGIA GIDGNGQSELIQALTGLRKAESGHIKLKGEDITNKKPRKITEHGVGHVPEDRHXYGLV LDMTLSSENIALQTYHQPKYSKNGMLNYSVINEHARELIEEYDVRTTNELVPAKALSGG NQOKAIIARIIVDRDPDLLIVANPTRGLDVGAIEFIHKRLIEQRDKYAVLLISFELEE ILNVSDRIAVIHEGEIVGIVDPKETSSEN</div>		

Further analysis of the NOV69a protein yielded the following properties shown in Table 69B.

Table 69B. Protein Sequence Properties NOV69a	
PSort analysis:	0.6000 probability located in nucleus; 0.1000 probability located in mitochondrial matrix space; 0.1000 probability located in lysosome (lumen); 0.0000 probability located in endoplasmic reticulum (membrane)
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV69a protein against the Geneseq database, a proprietary
 5 database that contains sequences published in patents and patent publication, yielded
 several homologous proteins shown in Table 69C.

Table 69C. Geneseq Results for NOV69a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV69a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAE16766	Human transporter and ion channel-3 (TRICH-3) protein - Homo sapiens, 485 aa. [WO200192304-A2, 06-DEC-2001]	8..492 1..485	485/485 (100%) 485/485 (100%)	0.0
AAU35015	Enterococcus faecalis cellular proliferation protein #302 - Enterococcus faecalis, 513 aa. [WO200170955-A2, 27-SEP-2001]	8..492 1..485	395/485 (81%) 450/485 (92%)	0.0
AAU35015	Enterococcus faecalis cellular proliferation protein #302 - Enterococcus faecalis, 513 aa. [WO200170955-A2, 27-SEP-2001]	8..492 1..485	395/485 (81%) 450/485 (92%)	0.0
AAU38034	Streptococcus pneumoniae cellular proliferation protein #463 - Streptococcus pneumoniae, 511 aa. [WO200170955-A2, 27-SEP-2001]	2..491 3..492	352/490 (71%) 424/490 (85%)	0.0
AAU37574	Streptococcus pneumoniae cellular proliferation protein #3 - Streptococcus pneumoniae, 511 aa. [WO200170955-A2, 27-SEP-2001]	2..491 3..492	352/490 (71%) 424/490 (85%)	0.0

In a BLAST search of public sequence databases, the NOV69a protein was found to have homology to the proteins shown in the BLASTP data in Table 69D.

Table 69D. Public BLASTP Results for NOV69a				
Protein Accession Number	Protein/Organism/Length	NOV69a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
AAK99552	ABC TRANSPORTER ATP-BINDING PROTEIN-RIBOSE/GALACTOSE TRANSPORT - Streptococcus pneumoniae R6, 511 aa.	2..491 3..492	352/490 (71%) 424/490 (85%)	0.0
Q97RG9	SUGAR ABC TRANSPORTER, ATP-BINDING PROTEIN - Streptococcus pneumoniae, 511 aa.	2..491 3..492	352/490 (71%) 424/490 (85%)	0.0
Q99ZH5	PUTATIVE SUGAR ABC TRANSPORTER (ATP-BINDING PROTEIN) - Streptococcus pyogenes, 510 aa.	1..491 1..491	347/491 (70%) 424/491 (85%)	0.0
AAL97795	PUTATIVE SUGAR ABC TRANSPORTER (ATP-BINDING PROTEIN) - Streptococcus pyogenes MGAS8232, 510 aa.	1..491 1..491	347/491 (70%) 424/491 (85%)	0.0
Q8Y7A1	HYPOTHETICAL PROTEIN LMO1389 - Listeria monocytogenes, 513 aa.	5..491 4..490	345/487 (70%) 407/487 (82%)	0.0

PFam analysis predicts that the NOV69a protein contains the domains shown in the Table 69E.

Table 69E. Domain Analysis of NOV69a			
Pfam Domain	NOV69a Match Region	Identities/ Similarities for the Matched Region	Expect Value
ABC_tran	31..217	73/201 (36%) 144/201 (72%)	5.8e-49
ABC_tran	284..478	59/207 (29%) 134/207 (65%)	7.4e-18

Example 70.

- 5 The NOV70 clone was analyzed, and the nucleotide and encoded polypeptide sequences are shown in Table 70A.

Table 70A. NOV70 Sequence Analysis			
	SEQ ID NO: 199	689 bp	
NOV70a, CG99963-01 DNA Sequence	AGCGGTGAGCCACCGCACCTGGCCTATTATGTTTGTGTTAAAGCAGGGTCTCGGCTT GCAAATGCAGCCACCGACGCCAGGAGCTCTGTACGATCAGCCATGGTTAACCCACCG TGTTTCGACATCGCTGTGACGGCAAGCCCTTGGGCCGCGGTGTCTTCGAGCCGTTTGA AGACAAGGTTCCAAAGGCAGCAGAAAACTTTCGTGCTCTGACAATGTAGAGAAAGGA TTTGTTTATAAAGGTTCTTGCTTTTACAGAATTATTCAGGGTTTATGTGTGACGGGT GTGACTTCACATGCCATAATGGCACTGGTGGCAAGTACACCTGCGGGGAGAAATTTGA TGACGAGAGCTTCGTCTTAAAGCATAACATCCTGGCATCTGTCCATGGCAAATGCT GGACCCAACACAAATGGTCCCAGTGTTTTCATCTGCACTGCCAAGACTGAGTGGCTGG ATGGCAATCCTGTGGTCTTTGGCAAGGTGAAAGAAGGCATGAATATTGTGGAGGCCAT GGGGCACTTTGGGTCGAGAATGGCAAGACCAGCAAGAAGATCACCATTGCTGACTGT GGACAACCTTAAACAAGTTTGGCTTTTGTATTATCTTAACCACCAGACCATTCTTTTC TAGCTCAGAAGAGCACCCCTTACCCCATTTGCTTGCAGTATTTTTTTTTT		
	ORF Start: ATG at 101	ORF Stop: TAA at 590	
	SEQ ID NO: 200	163 aa	MW at 17503.8kD
NOV70a, CG99963-01 Protein Sequence	MVNPTVFDIAVDGKPLGRVSFEPPADKVPKAAENFRALSNVEKGFYKGSFPHRIIPG FMCQGGDFTCHNGTGGKYTCGEKFDDSFVLKHTHPGILSMANAGPNTNGSQCFICTA KTEWLDGNPNVVFVKVKEGMNIVEAMGHFGSENGKTSKKITIADCGQL		
	SEQ ID NO: 201	596 bp	
NOV70b, CG99963-02 DNA Sequence	TTGTTTTGTTAAAGCAGGGTCTCGGCTTTGCAATGCAGCCACCGACGCCAGGAGCT CTGTACGATCAGCCATGGTTAACCCACCGTGTTCGACATCGCTGTGACGGCAAGCC CTTGGACCGCGTGTCTTCGAGCCGTTTGACAGACAAGGTTTCCAAAGGCAGCAAAAAAC TTTCGTGCTCTGAGCAATGTAGAGAAAGGATTGGTTATAAAGGTTCTGCTTTCACAT GAATTATTCAGGGTTTATGTGTGACGGTGGTACTTCACATGCCATAATGGCACTGG TGGCAAGTACACCTGCGGGGAGAAATTTGATGACGAGAGCTTCGTCTTAAAGCATAACA CATCCTGGCATCTTGTCCATGGCAATGCTGGACCCAACACAAATGGTCCCAGTGT TCATCTGCACTGCCAAGACTGAGTGGCTGGATGGCAATCCTGTGGTCTTTGGCAAGGT GAAAGAAGGCATGAATATTGTGGAGGCCATGGGGCACTTTGGGTCCGAGAATGGCAAG ACCGGCAAGAAGATCACCATTGCTGACTGTGGACAACCTCTAAACAAGTTTGGCTTTTGT TTTATCTTAACCACCA		
	ORF Start: ATG at 73	ORF Stop: TAA at 562	
	SEQ ID NO: 202	163 aa	MW at 17531.8kD
NOV70b, CG99963-02 Protein Sequence	MVNPTVFDIAVDGKPLDRVSFEPPADKVPKAAENFRALSNVEKGFYKGSFPHRIIPG FMCQGGDFTCHNGTGGKYTCGEKFDDSFVLKHTHPGILSMANAGPNTNGSQCFICTA KTEWLDGNPNVVFVKVKEGMNIVEAMGHFGSENGKTGKKITIADCGQL		

Sequence comparison of the above protein sequences yields the following sequence relationships shown in Table 70B.

Table 70B. Comparison of NOV70a against NOV70b.		
Protein Sequence	NOV70a Residues/ Match Residues	Identities/ Similarities for the Matched Region
NOV70b	1..163	161/163 (98%)
	1..163	161/163 (98%)

Further analysis of the NOV70a protein yielded the following properties shown in Table 70C.

Table 70C. Protein Sequence Properties NOV70a	
PSort analysis:	0.6000 probability located in plasma membrane; 0.5689 probability located in microbody (peroxisome); 0.4500 probability located in cytoplasm; 0.1000 probability located in mitochondrial matrix space
SignalP analysis:	No Known Signal Sequence Predicted

A search of the NOV70a protein against the Geneseq database, a proprietary database that contains sequences published in patents and patent publication, yielded several homologous proteins shown in Table 70D.

Table 70D. Geneseq Results for NOV70a				
Geneseq Identifier	Protein/Organism/Length [Patent #, Date]	NOV70a Residues/ Match Residues	Identities/ Similarities for the Matched Region	Expect Value
AAU01195	Human cyclophilin A protein - Homo sapiens, 165 aa. [WO200132876-A2, 10-MAY-2001]	1..163 1..164	144/164 (87%) 148/164 (89%)	1e-79
AAW56028	Calcineurin protein - Mammalia, 165 aa. [WO9808956-A2, 05-MAR-1998]	1..163 1..164	144/164 (87%) 148/164 (89%)	1e-79
AAR13726	Bovine cyclophilin - Bos taurus, 163 aa. [US5047512-A, 10-SEP-1991]	2..163 1..163	143/163 (87%) 147/163 (89%)	2e-79
AAG65275	Haematopoietic stem cell proliferation agent related human protein #2 - Homo sapiens, 164 aa. [JP2001163798-A, 19-JUN-2001]	2..163 1..163	143/163 (87%) 147/163 (89%)	4e-79
AAP90431	Cyclophilin - Homo sapiens (human), 164 aa. [EP326067-A, 02-AUG-1989]	2..163 1..163	143/163 (87%) 147/163 (89%)	4e-79

In a BLAST search of public sequence databases, the NOV70a protein was found to have homology to the proteins shown in the BLASTP data in Table 70E.

Table 70E. Public BLASTP Results for NOV70a				
Protein Accession Number	Protein/Organism/Length	NOV70a Residues/ Match Residues	Identities/ Similarities for the Matched Portion	Expect Value
CAC39529	SEQUENCE 26 FROM PATENT WO0132876 - Homo sapiens (Human), 165 aa.	1..163 1..164	144/164 (87%) 148/164 (89%)	2e-79
Q9TTC6	CYCLOPHILIN 18 - Oryctolagus cuniculus (Rabbit), 164 aa.	1..163 1..164	143/164 (87%) 150/164 (91%)	4e-79
Q9BRU4	PEPTIDYLPROLYL ISOMERASE A (CYCLOPHILIN A) - Homo sapiens (Human), 165 aa.	1..163 1..164	144/164 (87%) 147/164 (88%)	5e-79
P04374	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Bos taurus (Bovine), and, 163 aa.	2..163 1..163	143/163 (87%) 147/163 (89%)	5e-79
P05092	Peptidyl-prolyl cis-trans isomerase A (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin A) (Cyclosporin A-binding protein) - Homo sapiens (Human), 164 aa.	2..163 1..163	143/163 (87%) 147/163 (89%)	9e-79

PFam analysis predicts that the NOV70a protein contains the domains shown in the Table 70F.

Table 70F. Domain Analysis of NOV70a			
Pfam Domain	NOV70a Match Region	Identities/ Similarities for the Matched Region	Expect Value
pro_isomerase	7..163	108/179 (60%) 140/179 (78%)	3.7e-87

5

Example B: Sequencing Methodology and Identification of NOVX Clones

1. **GeneCalling™ Technology:** This is a proprietary method of performing differential gene expression profiling between two or more samples developed at CuraGen and described by Shimkets, et al., "Gene expression analysis by transcript profiling

coupled to a gene database query" Nature Biotechnology 17:198-803 (1999). cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then digested with up to as many as 120 pairs of restriction enzymes and pairs of linker-adaptors specific for each pair of restriction enzymes were ligated to the appropriate end. The restriction digestion generates a mixture of unique cDNA gene fragments. Limited PCR amplification is performed with primers homologous to the linker adapter sequence where one primer is biotinylated and the other is fluorescently labeled. The doubly labeled material is isolated and the fluorescently labeled single strand is resolved by capillary gel electrophoresis. A computer algorithm compares the electropherograms from an experimental and control group for each of the restriction digestions. This and additional sequence-derived information is used to predict the identity of each differentially expressed gene fragment using a variety of genetic databases. The identity of the gene fragment is confirmed by additional, gene-specific competitive PCR or by isolation and sequencing of the gene fragment.

2. **SeqCalling™ Technology:** cDNA was derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then sequenced using CuraGen's proprietary SeqCalling technology. Sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled together, sometimes including public human sequences, using bioinformatic programs to produce a consensus sequence for each assembly. Each assembly is included in CuraGen Corporation's database. Sequences were included as components for assembly when the extent of identity with another component was at least 95% over 50 bp. Each assembly represents a gene or portion thereof and includes information on variants, such

as splice forms single nucleotide polymorphisms (SNPs), insertions, deletions and other sequence variations.

3. PathCalling™ Technology:

5 The NOVX nucleic acid sequences are derived by laboratory screening of cDNA library by the two-hybrid approach. cDNA fragments covering either the full length of the DNA sequence, or part of the sequence, or both, are sequenced. In silico prediction was based on sequences available in CuraGen Corporation's proprietary sequence databases or in the public human sequence databases, and provided either the full length DNA
10 sequence, or some portion thereof.

The laboratory screening was performed using the methods summarized below:

cDNA libraries were derived from various human samples representing multiple tissue types, normal and diseased states, physiological states, and developmental states from different donors. Samples were obtained as whole tissue, primary cells or tissue
15 cultured primary cells or cell lines. Cells and cell lines may have been treated with biological or chemical agents that regulate gene expression, for example, growth factors, chemokines or steroids. The cDNA thus derived was then directionally cloned into the appropriate two-hybrid vector (Gal4-activation domain (Gal4-AD) fusion). Such cDNA libraries as well as commercially available cDNA libraries from Clontech (Palo Alto, CA)
20 were then transferred from E.coli into a CuraGen Corporation proprietary yeast strain (disclosed in U. S. Patents 6,057,101 and 6,083,693, incorporated herein by reference in their entireties).

Gal4-binding domain (Gal4-BD) fusions of a CuraGen Corporation proprietary library of human sequences was used to screen multiple Gal4-AD fusion cDNA libraries
25 resulting in the selection of yeast hybrid diploids in each of which the Gal4-AD fusion contains an individual cDNA. Each sample was amplified using the polymerase chain reaction (PCR) using non-specific primers at the cDNA insert boundaries. Such PCR product was sequenced; sequence traces were evaluated manually and edited for corrections if appropriate. cDNA sequences from all samples were assembled together,
30 sometimes including public human sequences, using bioinformatic programs to produce a

consensus sequence for each assembly. Each assembly is included in CuraGen Corporation's database. Sequences were included as components for assembly when the extent of identity with another component was at least 95% over 50 bp. Each assembly represents a gene or portion thereof and includes information on variants, such as splice
5 forms single nucleotide polymorphisms (SNPs), insertions, deletions and other sequence variations.

Physical clone: the cDNA fragment derived by the screening procedure, covering the entire open reading frame is, as a recombinant DNA, cloned into pACT2 plasmid (Clontech) used to make the cDNA library. The recombinant plasmid is inserted into the
10 host and selected by the yeast hybrid diploid generated during the screening procedure by the mating of both CuraGen Corporation proprietary yeast strains N106' and YULH (U. S. Patents 6,057,101 and 6,083,693).

4. **RACE:** Techniques based on the polymerase chain reaction such as rapid
15 amplification of cDNA ends (RACE), were used to isolate or complete the predicted sequence of the cDNA of the invention. Usually multiple clones were sequenced from one or more human samples to derive the sequences for fragments. Various human tissue samples from different donors were used for the RACE reaction. The sequences derived from these procedures were included in the SeqCalling Assembly process described in
20 preceding paragraphs.

5. **Exon Linking:** The NOVX target sequences identified in the present invention were subjected to the exon linking process to confirm the sequence. PCR primers were designed by starting at the most upstream sequence available, for the forward primer, and
25 at the most downstream sequence available for the reverse primer. In each case, the sequence was examined, walking inward from the respective termini toward the coding sequence, until a suitable sequence that is either unique or highly selective was encountered, or, in the case of the reverse primer, until the stop codon was reached. Such primers were designed based on in silico predictions for the full length cDNA, part (one or
30 more exons) of the DNA or protein sequence of the target sequence, or by translated homology of the predicted exons to closely related human sequences from other species. These primers were then employed in PCR amplification based on the following pool of

human cDNAs: adrenal gland, bone marrow, brain - amygdala, brain - cerebellum, brain - hippocampus, brain - substantia nigra, brain - thalamus, brain - whole, fetal brain, fetal kidney, fetal liver, fetal lung, heart, kidney, lymphoma - Raji, mammary gland, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord, spleen, stomach, testis, thyroid, trachea, uterus. Usually the resulting amplicons were gel purified, cloned and sequenced to high redundancy. The PCR product derived from exon linking was cloned into the pCR2.1 vector from Invitrogen. The resulting bacterial clone has an insert covering the entire open reading frame cloned into the pCR2.1 vector. The resulting sequences from all clones were assembled with themselves, with other fragments in CuraGen Corporation's database and with public ESTs. Fragments and ESTs were included as components for an assembly when the extent of their identity with another component of the assembly was at least 95% over 50 bp. In addition, sequence traces were evaluated manually and edited for corrections if appropriate. These procedures provide the sequence reported herein.

6. Physical Clone:

Exons were predicted by homology and the intron/exon boundaries were determined using standard genetic rules. Exons were further selected and refined by means of similarity determination using multiple BLAST (for example, tBlastN, BlastX, and BlastN) searches, and, in some instances, GeneScan and Grail. Expressed sequences from both public and proprietary databases were also added when available to further define and complete the gene sequence. The DNA sequence was then manually corrected for apparent inconsistencies thereby obtaining the sequences encoding the full-length protein.

The PCR product derived by exon linking, covering the entire open reading frame, was cloned into the pCR2.1 vector from Invitrogen to provide clones used for expression and screening purposes.

Example C: Quantitative expression analysis of clones in various cells and tissues

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR). RTQ PCR was performed on an Applied Biosystems ABI PRISM® 7700 or an ABI PRISM® 7900 HT Sequence

Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing normal tissues and cancer cell lines), Panel 2 (containing samples derived from tissues from normal and cancer sources), Panel 3 (containing cancer cell lines), Panel 4 (containing cells and cell lines from normal tissues and cells related to inflammatory conditions), Panel 5D/5I (containing human tissues and cell lines with an emphasis on metabolic diseases), AI_comprehensive_panel (containing normal tissue and samples from autoimmune diseases), Panel CNSD.01 (containing central nervous system samples from normal and diseased brains) and CNS_neurodegeneration_panel (containing samples from normal and Alzheimer's diseased brains).

RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase using probe and primer sets designed to amplify across the span of a single exon.

First, the RNA samples were normalized to reference nucleic acids such as constitutively expressed genes (for example, β -actin and GAPDH). Normalized RNA (5 μ l) was converted to cDNA and analyzed by RTQ-PCR using One Step RT-PCR Master Mix Reagents (Applied Biosystems; Catalog No. 4309169) and gene-specific primers according to the manufacturer's instructions.

In other cases, non-normalized RNA samples were converted to single strand cDNA (sscDNA) using Superscript II (Invitrogen Corporation; Catalog No. 18064-147) and random hexamers according to the manufacturer's instructions. Reactions containing up to 10 μ g of total RNA were performed in a volume of 20 μ l and incubated for 60 minutes at 42°C. This reaction can be scaled up to 50 μ g of total RNA in a final volume of 100 μ l. sscDNA samples are then normalized to reference nucleic acids as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions.

Probes and primers were designed for each assay according to Applied Biosystems Primer Express Software package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer

concentration = 250 nM, primer melting temperature (T_m) range = 58°-60°C, primer optimal T_m = 59°C, maximum primer difference = 2°C, probe does not have 5'G, probe T_m must be 10°C greater than primer T_m , amplicon size 75bp to 100bp. The probes and primers selected (see below) were synthesized by Synthesgen (Houston, TX, USA). Probes
5 were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900nM each, and probe, 200nM.

PCR conditions: When working with RNA samples, normalized RNA from each
10 tissue and each cell line was spotted in each well of either a 96 well or a 384-well PCR plate (Applied Biosystems). PCR cocktails included either a single gene specific probe and primers set, or two multiplexed probe and primers sets (a set specific for the target clone and another gene-specific set multiplexed with the target probe). PCR reactions were set up using TaqMan® One-Step RT-PCR Master Mix (Applied Biosystems, Catalog No.
15 4313803) following manufacturer's instructions. Reverse transcription was performed at 48°C for 30 minutes followed by amplification/PCR cycles as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were recorded as CT values (cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the
20 lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying by 100.

When working with sscDNA samples, normalized sscDNA was used as described previously for RNA samples. PCR reactions containing one or two sets of probe and
25 primers were set up as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions. PCR amplification was performed as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were analyzed and processed as described previously.

Panels 1, 1.1, 1.2, and 1.3D

30 The plates for Panels 1, 1.1, 1.2 and 1.3D include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in these panels are broken into 2 classes: samples derived from cultured cell lines

and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in these panels are widely available
 5 through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on these panels are comprised of samples derived from all major organ systems from single adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult
 10 kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose.

15 In the results for Panels 1, 1.1, 1.2 and 1.3D, the following abbreviations are used:
 ca. = carcinoma,
 * = established from metastasis,
 met = metastasis,
 s cell var = small cell variant,
 20 non-s = non-sm = non-small,
 squam = squamous,
 pl. eff = pl effusion = pleural effusion,
 glio = glioma,
 astro = astrocytoma, and
 25 neuro = neuroblastoma.

General_screening_panel_v1.4 and General_screening_panel_v1.5

The plates for Panels 1.4 and 1.5 include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples
 30 in Panels 1.4 and 1.5 are broken into 2 classes: samples derived from cultured cell lines and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in Panel 1.4 are widely available
 35 through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal

tissues found on Panels 1.4 and 1.5 are comprised of pools of samples derived from all major organ systems from 2 to 5 different adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various
5 regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose. Abbreviations are as described for Panels 1, 1.1, 1.2, and 1.3D.

Panels 2D and 2.2

10 The plates for Panels 2D and 2.2 generally include 2 control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by surgeons working in close cooperation with the National Cancer Institute's Cooperative Human Tissue Network (CHTN) or the National Disease Research Initiative (NDRI). The tissues are derived from human malignancies and in cases where indicated many malignant
15 tissues have "matched margins" obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologist at NDRI or CHTN). This analysis provides a gross histopathological assessment of tumor differentiation grade.
20 Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (i.e. immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly
25 people or sudden death victims (accidents, etc.). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen.

Panel 3D

30 The plates of Panel 3D are comprised of 94 cDNA samples and two control samples. Specifically, 92 of these samples are derived from cultured human cancer cell lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines

are generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas, ovarian/uterine/cervical, gastric, colon, lung and CNS cancer cell lines. In addition, there are two independent samples of cerebellum. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. The cell lines in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.

Panels 4D, 4R, and 4.1D

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4R) or cDNA (Panels 4D/4.1D) isolated from various human cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene, La Jolla, CA) and thymus and kidney (Clontech) was employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA). Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and/or 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-5ng/ml, TNF alpha at approximately 5-10ng/ml, IFN gamma at approximately 20-50ng/ml, IL-4 at approximately 5-10ng/ml, IL-9 at approximately 5-10ng/ml, IL-13 at approximately 5-10ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco/Life Technologies,

Rockville, MD), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20ng/ml PMA and 1-2 μ g/ml ionomycin, IL-12 at 5-10ng/ml, IFN gamma at 20-50ng/ml and IL-18 at 5-10ng/ml for 6 hours. In some cases, mononuclear cells were

5 cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5 μ g/ml. Samples were taken at 24, 48 and 72 hours for RNA preparation. MLR (mixed lymphocyte reaction) samples were obtained by taking blood from two

10 donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately 2×10^6 cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol (5.5×10^{-5} M) (Gibco), and 10mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

15 Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +ve VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum (FCS) (Hyclone, Logan, UT), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco),

20 50ng/ml GMCSF and 5ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10mM Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with

25 lipopolysaccharide (LPS) at 100ng/ml. Dendritic cells were also stimulated with anti-CD40 monoclonal antibody (Pharmingen) at 10 μ g/ml for 6 and 12-14 hours.

30 CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56, CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. CD45RO beads were then used to isolate the CD45RO CD4 lymphocytes with

the remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), and 10mM Hepes (Gibco) and plated at 10⁶cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5 μ g/ml anti-CD28 (Pharmingen) and 3 μ g/ml anti-CD3 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested the cells and expanded them in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), and 10mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), and 10mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resuspended at 10⁶cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), and 10mM Hepes (Gibco). To activate the cells, we used PWM at 5 μ g/ml or anti-CD40 (Pharmingen) at approximately 10 μ g/ml and IL-4 at 5-10ng/ml. Cells were harvested for RNA preparation at 24,48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon plates were coated overnight with 10 μ g/ml anti-CD28 (Pharmingen) and 2 μ g/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at 10⁵-10⁶cells/ml in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), 10mM Hepes (Gibco) and IL-2 (4ng/ml). IL-12 (5ng/ml) and anti-IL4 (1 μ g/ml) were used to direct to Th1, while IL-4 (5ng/ml) and anti-IFN gamma (1 μ g/ml) were used to direct to Th2 and IL-10 at 5ng/ml was used to

direct to Tr1. After 4-5 days, the activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), 10mM Hepes (Gibco) and IL-2 (1ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1 μ g/ml) to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours following the second and third activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1mM dbcAMP at 5x10⁵cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to 5x10⁵cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), 10mM Hepes (Gibco). RNA was either prepared from resting cells or cells activated with PMA at 10ng/ml and ionomycin at 1 μ g/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5x10⁻⁵M (Gibco), and 10mM Hepes (Gibco). CCD1106 cells were activated for 6 and 14 hours with approximately 5 ng/ml TNF alpha and 1ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5ng/ml IL-4, 5ng/ml IL-9, 5ng/ml IL-13 and 25ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately 10⁷cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15ml Falcon Tube. An equal volume of

isopropanol was added and left at -20°C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet was redissolved in 300µl of RNase-free water and 35µl buffer (Promega) 5µl DTT, 7µl RNasin and 8µl DNase were added. The tube was incubated at 37°C for 30 minutes to
5 remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNase free water. RNA was stored at -80°C.

AI_comprehensive panel_v1.0

The plates for AI_comprehensive panel_v1.0 include two control wells and 89 test
10 samples comprised of cDNA isolated from surgical and postmortem human tissues obtained from the Backus Hospital and Clinomics (Frederick, MD). Total RNA was extracted from tissue samples from the Backus Hospital in the Facility at CuraGen. Total RNA from other tissues was obtained from Clinomics.

Joint tissues including synovial fluid, synovium, bone and cartilage were obtained
15 from patients undergoing total knee or hip replacement surgery at the Backus Hospital. Tissue samples were immediately snap frozen in liquid nitrogen to ensure that isolated RNA was of optimal quality and not degraded. Additional samples of osteoarthritis and rheumatoid arthritis joint tissues were obtained from Clinomics. Normal control tissues were supplied by Clinomics and were obtained during autopsy of trauma victims.

20 Surgical specimens of psoriatic tissues and adjacent matched tissues were provided as total RNA by Clinomics. Two male and two female patients were selected between the ages of 25 and 47. None of the patients were taking prescription drugs at the time samples were isolated.

Surgical specimens of diseased colon from patients with ulcerative colitis and
25 Crohns disease and adjacent matched tissues were obtained from Clinomics. Bowel tissue from three female and three male Crohn's patients between the ages of 41-69 were used. Two patients were not on prescription medication while the others were taking dexamethasone, phenobarbital, or tylenol. Ulcerative colitis tissue was from three male and four female patients. Four of the patients were taking lebid and two were on
30 phenobarbital.

Total RNA from post mortem lung tissue from trauma victims with no disease or with emphysema, asthma or COPD was purchased from Clinomics. Emphysema patients

ranged in age from 40-70 and all were smokers, this age range was chosen to focus on patients with cigarette-linked emphysema and to avoid those patients with alpha-1 anti-trypsin deficiencies. Asthma patients ranged in age from 36-75, and excluded smokers to prevent those patients that could also have COPD. COPD patients ranged in age from 35-80 and included both smokers and non-smokers. Most patients were taking corticosteroids, and bronchodilators.

In the labels employed to identify tissues in the AI_comprehensive panel_v1.0 panel, the following abbreviations are used:

AI = Autoimmunity
Syn = Synovial
Normal = No apparent disease
Rep22 /Rep20 = individual patients
RA = Rheumatoid arthritis
Backus = From Backus Hospital
OA = Osteoarthritis
(SS) (BA) (MF) = Individual patients
Adj = Adjacent tissue
Match control = adjacent tissues
-M = Male
-F = Female
COPD = Chronic obstructive pulmonary disease

Panels 5D and 5I

The plates for Panel 5D and 5I include two control wells and a variety of cDNAs isolated from human tissues and cell lines with an emphasis on metabolic diseases. Metabolic tissues were obtained from patients enrolled in the Gestational Diabetes study. Cells were obtained during different stages in the differentiation of adipocytes from human mesenchymal stem cells. Human pancreatic islets were also obtained.

In the Gestational Diabetes study subjects are young (18 - 40 years), otherwise healthy women with and without gestational diabetes undergoing routine (elective) Caesarean section. After delivery of the infant, when the surgical incisions were being repaired/closed, the obstetrician removed a small sample (<1 cc) of the exposed metabolic tissues during the closure of each surgical level. The biopsy material was rinsed in sterile saline, blotted and fast frozen within 5 minutes from the time of removal. The tissue was then flash frozen in liquid nitrogen and stored, individually, in sterile screw-top tubes and kept on dry ice for shipment to or to be picked up by CuraGen. The metabolic tissues of

interest include uterine wall (smooth muscle), visceral adipose, skeletal muscle (rectus) and subcutaneous adipose. Patient descriptions are as follows:

- Patient 2: Diabetic Hispanic, overweight, not on insulin
- Patient 7-9: Nondiabetic Caucasian and obese (BMI>30)
- Patient 10: Diabetic Hispanic, overweight, on insulin
- Patient 11: Nondiabetic African American and overweight
- Patient 12: Diabetic Hispanic on insulin

Adipocyte differentiation was induced in donor progenitor cells obtained from Osirus (a division of Clonetics/BioWhittaker) in triplicate, except for Donor 3U which had only two replicates. Scientists at Clonetics isolated, grew and differentiated human mesenchymal stem cells (HuMSCs) for CuraGen based on the published protocol found in Mark F. Pittenger, et al., Multilineage Potential of Adult Human Mesenchymal Stem Cells Science Apr 2 1999: 143-147. Clonetics provided Trizol lysates or frozen pellets suitable for mRNA isolation and ds cDNA production. A general description of each donor is as follows:

- Donor 2 and 3 U: Mesenchymal Stem cells, Undifferentiated Adipose
 - Donor 2 and 3 AM: Adipose, AdiposeMidway Differentiated
 - Donor 2 and 3 AD: Adipose, Adipose Differentiated
- Human cell lines were generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: kidney proximal convoluted tubule, uterine smooth muscle cells, small intestine, liver HepG2 cancer cells, heart primary stromal cells, and adrenal cortical adenoma cells. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. All samples were processed at CuraGen to produce single stranded cDNA.

Panel 5I contains all samples previously described with the addition of pancreatic islets from a 58 year old female patient obtained from the Diabetes Research Institute at the University of Miami School of Medicine. Islet tissue was processed to total RNA at an outside source and delivered to CuraGen for addition to panel 5I.

In the labels employed to identify tissues in the 5D and 5I panels, the following abbreviations are used:

- GO Adipose = Greater Omentum Adipose
- SK = Skeletal Muscle
- UT = Uterus
- PL = Placenta
- AD = Adipose Differentiated

AM = Adipose Midway Differentiated
U = Undifferentiated Stem Cells

Panel CNSD.01

5 The plates for Panel CNSD.01 include two control wells and 94 test samples
comprised of cDNA isolated from postmortem human brain tissue obtained from the
Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors
between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in
liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to
10 confirm diagnoses with clear associated neuropathology.

 Disease diagnoses are taken from patient records. The panel contains two brains
from each of the following diagnoses: Alzheimer's disease, Parkinson's disease,
Huntington's disease, Progressive Supranuclear Palsy, Depression, and "Normal controls".
Within each of these brains, the following regions are represented: cingulate gyrus,
15 temporal pole, globus palladus, substantia nigra, Brodman Area 4 (primary motor strip),
Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area
17 (occipital cortex). Not all brain regions are represented in all cases; e.g., Huntington's
disease is characterized in part by neurodegeneration in the globus palladus, thus this
region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's
20 disease is characterized by degeneration of the substantia nigra making this region more
difficult to obtain. Normal control brains were examined for neuropathology and found to
be free of any pathology consistent with neurodegeneration.

 In the labels employed to identify tissues in the CNS panel, the following
abbreviations are used:

25 PSP = Progressive supranuclear palsy
 Sub Nigra = Substantia nigra
 Glob Palladus= Globus palladus
 Temp Pole = Temporal pole
 Cing Gyr = Cingulate gyrus
30 BA 4 = Brodman Area 4

Panel CNS_Neurodegeneration_V1.0

 The plates for Panel CNS_Neurodegeneration_V1.0 include two control wells and
47 test samples comprised of cDNA isolated from postmortem human brain tissue
35 obtained from the Harvard Brain Tissue Resource Center (McLean Hospital) and the

Human Brain and Spinal Fluid Resource Center (VA Greater Los Angeles Healthcare System). Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains six brains from Alzheimer's disease (AD) patients, and eight brains from "Normal controls" who showed no evidence of dementia prior to death. The eight normal control brains are divided into two categories: Controls with no dementia and no Alzheimer's like pathology (Controls) and controls with no dementia but evidence of severe Alzheimer's like pathology, (specifically senile plaque load rated as level 3 on a scale of 0-3; 0 = no evidence of plaques, 3 = severe AD senile plaque load). Within each of these brains, the following regions are represented: hippocampus, temporal cortex (Brodmann Area 21), parietal cortex (Brodmann area 7), and occipital cortex (Brodmann area 17). These regions were chosen to encompass all levels of neurodegeneration in AD. The hippocampus is a region of early and severe neuronal loss in AD; the temporal cortex is known to show neurodegeneration in AD after the hippocampus; the parietal cortex shows moderate neuronal death in the late stages of the disease; the occipital cortex is spared in AD and therefore acts as a "control" region within AD patients. Not all brain regions are represented in all cases.

In the labels employed to identify tissues in the CNS_Neurodegeneration_V1.0 panel, the following abbreviations are used:

AD = Alzheimer's disease brain; patient was demented and showed AD-like pathology upon autopsy
Control = Control brains; patient not demented, showing no neuropathology
Control (Path) = Control brains; patient not demented but showing severe AD-like pathology
SupTemporal Ctx = Superior Temporal Cortex
Inf Temporal Ctx = Inferior Temporal Cortex

The quantitative expression of various clones was assessed using microtiter plates containing RNA samples from a variety of normal and pathology-derived cells, cell lines and tissues using real time quantitative PCR (RTQ PCR). RTQ PCR was performed on an Applied Biosystems ABI PRISM® 7700 or an ABI PRISM® 7900 HT Sequence

Detection System. Various collections of samples are assembled on the plates, and referred to as Panel 1 (containing normal tissues and cancer cell lines), Panel 2 (containing samples derived from tissues from normal and cancer sources), Panel 3 (containing cancer cell lines), Panel 4 (containing cells and cell lines from normal tissues and cells related to inflammatory conditions), Panel 5D/5I (containing human tissues and cell lines with an emphasis on metabolic diseases), AI_comprehensive_panel (containing normal tissue and samples from autoinflammatory diseases), Panel CNSD.01 (containing samples from normal and diseased brains) and CNS_neurodegeneration_panel (containing samples from normal and Alzheimer's diseased brains).

10 RNA integrity from all samples is controlled for quality by visual assessment of agarose gel electropherograms using 28S and 18S ribosomal RNA staining intensity ratio as a guide (2:1 to 2.5:1 28s:18s) and the absence of low molecular weight RNAs that would be indicative of degradation products. Samples are controlled against genomic DNA contamination by RTQ PCR reactions run in the absence of reverse transcriptase
15 using probe and primer sets designed to amplify across the span of a single exon.

First, the RNA samples were normalized to reference nucleic acids such as constitutively expressed genes (for example, β -actin and GAPDH). Normalized RNA (5 μ l) was converted to cDNA and analyzed by RTQ-PCR using One Step RT-PCR Master Mix Reagents (Applied Biosystems; Catalog No. 4309169) and gene-specific primers
20 according to the manufacturer's instructions.

In other cases, non-normalized RNA samples were converted to single strand cDNA (sscDNA) using Superscript II (Invitrogen Corporation; Catalog No. 18064-147) and random hexamers according to the manufacturer's instructions. Reactions containing up to 10 μ g of total RNA were performed in a volume of 20 μ l and incubated for 60
25 minutes at 42°C. This reaction can be scaled up to 50 μ g of total RNA in a final volume of 100 μ l. sscDNA samples are then normalized to reference nucleic acids as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions.

Probes and primers were designed for each assay according to Applied Biosystems
30 Primer Express Software package (version I for Apple Computer's Macintosh Power PC) or a similar algorithm using the target sequence as input. Default settings were used for reaction conditions and the following parameters were set before selecting primers: primer

concentration = 250 nM, primer melting temperature (T_m) range = 58°-60°C, primer optimal T_m = 59°C, maximum primer difference = 2°C, probe does not have 5'G, probe T_m must be 10°C greater than primer T_m , amplicon size 75bp to 100bp. The probes and primers selected (see below) were synthesized by SyntheGen (Houston, TX, USA). Probes were double purified by HPLC to remove uncoupled dye and evaluated by mass spectroscopy to verify coupling of reporter and quencher dyes to the 5' and 3' ends of the probe, respectively. Their final concentrations were: forward and reverse primers, 900nM each, and probe, 200nM.

PCR conditions: When working with RNA samples, normalized RNA from each tissue and each cell line was spotted in each well of either a 96 well or a 384-well PCR plate (Applied Biosystems). PCR cocktails included either a single gene specific probe and primers set, or two multiplexed probe and primers sets (a set specific for the target clone and another gene-specific set multiplexed with the target probe). PCR reactions were set up using TaqMan® One-Step RT-PCR Master Mix (Applied Biosystems, Catalog No. 4313803) following manufacturer's instructions. Reverse transcription was performed at 48°C for 30 minutes followed by amplification/PCR cycles as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were recorded as CT values (cycle at which a given sample crosses a threshold level of fluorescence) using a log scale, with the difference in RNA concentration between a given sample and the sample with the lowest CT value being represented as 2 to the power of delta CT. The percent relative expression is then obtained by taking the reciprocal of this RNA difference and multiplying by 100.

When working with sscDNA samples, normalized sscDNA was used as described previously for RNA samples. PCR reactions containing one or two sets of probe and primers were set up as described previously, using 1X TaqMan® Universal Master mix (Applied Biosystems; catalog No. 4324020), following the manufacturer's instructions. PCR amplification was performed as follows: 95°C 10 min, then 40 cycles of 95°C for 15 seconds, 60°C for 1 minute. Results were analyzed and processed as described previously.

Panels 1, 1.1, 1.2, and 1.3D

The plates for Panels 1, 1.1, 1.2 and 1.3D include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in these panels are broken into 2 classes: samples derived from cultured cell lines

and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer, gastric cancer and pancreatic cancer. Cell lines used in these panels are widely available through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on these panels are comprised of samples derived from all major organ systems from single adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon, bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose.

In the results for Panels 1, 1.1, 1.2 and 1.3D, the following abbreviations are used:

ca. = carcinoma,
* = established from metastasis,
met = metastasis,
s cell var = small cell variant,
non-s = non-sm = non-small,
squam = squamous,
pl. eff = pl effusion = pleural effusion,
glio = glioma,
astro = astrocytoma, and
neuro = neuroblastoma.

General_screening_panel_v1.4 and General_screening_panel_v1.5

The plates for Panels 1.4 and 1.5 include 2 control wells (genomic DNA control and chemistry control) and 94 wells containing cDNA from various samples. The samples in Panels 1.4 and 1.5 are broken into 2 classes: samples derived from cultured cell lines and samples derived from primary normal tissues. The cell lines are derived from cancers of the following types: lung cancer, breast cancer, melanoma, colon cancer, prostate cancer, CNS cancer, squamous cell carcinoma, ovarian cancer, liver cancer, renal cancer,

gastric cancer and pancreatic cancer. Cell lines used in Panels 1.4 and 1.5 are widely available through the American Type Culture Collection (ATCC), a repository for cultured cell lines, and were cultured using the conditions recommended by the ATCC. The normal tissues found on Panels 1.4 and 1.5 are comprised of pools of samples derived from all
5 major organ systems from 2 to 5 different adult individuals or fetuses. These samples are derived from the following organs: adult skeletal muscle, fetal skeletal muscle, adult heart, fetal heart, adult kidney, fetal kidney, adult liver, fetal liver, adult lung, fetal lung, various regions of the brain, the spleen, bone marrow, lymph node, pancreas, salivary gland, pituitary gland, adrenal gland, spinal cord, thymus, stomach, small intestine, colon,
10 bladder, trachea, breast, ovary, uterus, placenta, prostate, testis and adipose. Abbreviations are as described for Panels 1, 1.1, 1.2, and 1.3D.

Panels 2D, 2.2, 2.3 and 2.4

The plates for Panels 2D, 2.2, 2.3 and 2.4 generally include 2 control wells and 94 test samples composed of RNA or cDNA isolated from human tissue procured by
15 surgeons working in close cooperation with the National Cancer Institute's Cooperative Human Tissue Network (CHTN) or the National Disease Research Initiative (NDRI) or from Ardaïs or Clinomics). The tissues are derived from human malignancies and in cases where indicated many malignant tissues have "matched margins" obtained from noncancerous tissue just adjacent to the tumor. These are termed normal adjacent tissues
20 and are denoted "NAT" in the results below. The tumor tissue and the "matched margins" are evaluated by two independent pathologists (the surgical pathologists and again by a pathologist at NDRI/ CHTN/Ardaïs/Clinomics). Unmatched RNA samples from tissues without malignancy (normal tissues) were also obtained from Ardaïs or Clinomics. This analysis provides a gross histopathological assessment of tumor differentiation grade.
25 Moreover, most samples include the original surgical pathology report that provides information regarding the clinical stage of the patient. These matched margins are taken from the tissue surrounding (i.e. immediately proximal) to the zone of surgery (designated "NAT", for normal adjacent tissue, in Table RR). In addition, RNA and cDNA samples were obtained from various human tissues derived from autopsies performed on elderly
30 people or sudden death victims (accidents, etc.). These tissues were ascertained to be free of disease and were purchased from various commercial sources such as Clontech (Palo Alto, CA), Research Genetics, and Invitrogen.

HASS Panel v 1.0

The HASS panel v 1.0 plates are comprised of 93 cDNA samples and two controls. Specifically, 81 of these samples are derived from cultured human cancer cell lines that had been subjected to serum starvation, acidosis and anoxia for different time periods as well as controls for these treatments, 3 samples of human primary cells, 9 samples of malignant brain cancer (4 medulloblastomas and 5 glioblastomas) and 2 controls. The human cancer cell lines are obtained from ATCC (American Type Culture Collection) and fall into the following tissue groups: breast cancer, prostate cancer, bladder carcinomas, pancreatic cancers and CNS cancer cell lines. These cancer cells are all cultured under standard recommended conditions. The treatments used (serum starvation, acidosis and anoxia) have been previously published in the scientific literature. The primary human cells were obtained from Clonetics (Walkersville, MD) and were grown in the media and conditions recommended by Clonetics. The malignant brain cancer samples are obtained as part of a collaboration (Henry Ford Cancer Center) and are evaluated by a pathologist prior to CuraGen receiving the samples. RNA was prepared from these samples using the standard procedures. The genomic and chemistry control wells have been described previously.

Panel 3D

The plates of Panel 3D are comprised of 94 cDNA samples and two control samples. Specifically, 92 of these samples are derived from cultured human cancer cell lines, 2 samples of human primary cerebellar tissue and 2 controls. The human cell lines are generally obtained from ATCC (American Type Culture Collection), NCI or the German tumor cell bank and fall into the following tissue groups: Squamous cell carcinoma of the tongue, breast cancer, prostate cancer, melanoma, epidermoid carcinoma, sarcomas, bladder carcinomas, pancreatic cancers, kidney cancers, leukemias/lymphomas, ovarian/uterine/cervical, gastric, colon, lung and CNS cancer cell lines. In addition, there are two independent samples of cerebellum. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. The cell lines in panel 3D and 1.3D are of the most common cell lines used in the scientific literature.

Panels 4D, 4R, and 4.1D

Panel 4 includes samples on a 96 well plate (2 control wells, 94 test samples) composed of RNA (Panel 4R) or cDNA (Panels 4D/4.1D) isolated from various human

cell lines or tissues related to inflammatory conditions. Total RNA from control normal tissues such as colon and lung (Stratagene, La Jolla, CA) and thymus and kidney (Clontech) was employed. Total RNA from liver tissue from cirrhosis patients and kidney from lupus patients was obtained from BioChain (Biochain Institute, Inc., Hayward, CA).

- 5 Intestinal tissue for RNA preparation from patients diagnosed as having Crohn's disease and ulcerative colitis was obtained from the National Disease Research Interchange (NDRI) (Philadelphia, PA).

Astrocytes, lung fibroblasts, dermal fibroblasts, coronary artery smooth muscle cells, small airway epithelium, bronchial epithelium, microvascular dermal endothelial
10 cells, microvascular lung endothelial cells, human pulmonary aortic endothelial cells, human umbilical vein endothelial cells were all purchased from Clonetics (Walkersville, MD) and grown in the media supplied for these cell types by Clonetics. These primary cell types were activated with various cytokines or combinations of cytokines for 6 and/or 12-14 hours, as indicated. The following cytokines were used; IL-1 beta at approximately 1-
15 5ng/ml, TNF alpha at approximately 5-10ng/ml, IFN gamma at approximately 20-50ng/ml, IL-4 at approximately 5-10ng/ml, IL-9 at approximately 5-10ng/ml, IL-13 at approximately 5-10ng/ml. Endothelial cells were sometimes starved for various times by culture in the basal media from Clonetics with 0.1% serum.

Mononuclear cells were prepared from blood of employees at CuraGen
20 Corporation, using Ficoll. LAK cells were prepared from these cells by culture in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco/Life Technologies, Rockville, MD), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) and Interleukin 2 for 4-6 days. Cells were then either activated with 10-20ng/ml PMA and 1-2μg/ml ionomycin, IL-12 at 5-10ng/ml, IFN gamma at 20-
25 50ng/ml and IL-18 at 5-10ng/ml for 6 hours. In some cases, mononuclear cells were cultured for 4-5 days in DMEM 5% FCS (Hyclone), 100μM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) with PHA (phytohemagglutinin) or PWM (pokeweed mitogen) at approximately 5μg/ml. Samples were taken at 24, 48 and 72 hours for RNA preparation.
30 MLR (mixed lymphocyte reaction) samples were obtained by taking blood from two donors, isolating the mononuclear cells using Ficoll and mixing the isolated mononuclear cells 1:1 at a final concentration of approximately 2×10^6 cells/ml in DMEM 5% FCS

(Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol (5.5×10^{-5} M) (Gibco), and 10mM Hepes (Gibco). The MLR was cultured and samples taken at various time points ranging from 1- 7 days for RNA preparation.

Monocytes were isolated from mononuclear cells using CD14 Miltenyi Beads, +ve
5 VS selection columns and a Vario Magnet according to the manufacturer's instructions. Monocytes were differentiated into dendritic cells by culture in DMEM 5% fetal calf serum (FCS) (Hyclone, Logan, UT), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco), 50ng/ml GMCSF and 5ng/ml IL-4 for 5-7 days. Macrophages were prepared by culture of
10 monocytes for 5-7 days in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10mM Hepes (Gibco) and 10% AB Human Serum or MCSF at approximately 50ng/ml. Monocytes, macrophages and dendritic cells were stimulated for 6 and 12-14 hours with lipopolysaccharide (LPS) at 100ng/ml. Dendritic cells were also stimulated with anti-
15 CD40 monoclonal antibody (Pharmingen) at 10 μ g/ml for 6 and 12-14 hours.

CD4 lymphocytes, CD8 lymphocytes and NK cells were also isolated from mononuclear cells using CD4, CD8 and CD56 Miltenyi beads, positive VS selection columns and a Vario Magnet according to the manufacturer's instructions. CD45RA and CD45RO CD4 lymphocytes were isolated by depleting mononuclear cells of CD8, CD56,
20 CD14 and CD19 cells using CD8, CD56, CD14 and CD19 Miltenyi beads and positive selection. CD45RO beads were then used to isolate the CD45RO CD4 lymphocytes with the remaining cells being CD45RA CD4 lymphocytes. CD45RA CD4, CD45RO CD4 and CD8 lymphocytes were placed in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and
25 10mM Hepes (Gibco) and plated at 10^6 cells/ml onto Falcon 6 well tissue culture plates that had been coated overnight with 0.5 μ g/ml anti-CD28 (Pharmingen) and 3 μ g/ml anti-CD3 (OKT3, ATCC) in PBS. After 6 and 24 hours, the cells were harvested for RNA preparation. To prepare chronically activated CD8 lymphocytes, we activated the isolated CD8 lymphocytes for 4 days on anti-CD28 and anti-CD3 coated plates and then harvested
30 the cells and expanded them in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) and IL-2. The expanded CD8 cells were then activated again with

plate bound anti-CD3 and anti-CD28 for 4 days and expanded as before. RNA was isolated 6 and 24 hours after the second activation and after 4 days of the second expansion culture. The isolated NK cells were cultured in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco),
5 mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco) and IL-2 for 4-6 days before RNA was prepared.

To obtain B cells, tonsils were procured from NDRI. The tonsil was cut up with sterile dissecting scissors and then passed through a sieve. Tonsil cells were then spun down and resuspended at 10^6 cells/ml in DMEM 5% FCS (Hyclone), 100µM non essential
10 amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco). To activate the cells, we used PWM at 5µg/ml or anti-CD40 (Pharmingen) at approximately 10µg/ml and IL-4 at 5-10ng/ml. Cells were harvested for RNA preparation at 24, 48 and 72 hours.

To prepare the primary and secondary Th1/Th2 and Tr1 cells, six-well Falcon
15 plates were coated overnight with 10µg/ml anti-CD28 (Pharmingen) and 2µg/ml OKT3 (ATCC), and then washed twice with PBS. Umbilical cord blood CD4 lymphocytes (Poietic Systems, German Town, MD) were cultured at 10^5 - 10^6 cells/ml in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10mM Hepes (Gibco) and IL-2 (4ng/ml).
20 IL-12 (5ng/ml) and anti-IL4 (1µg/ml) were used to direct to Th1, while IL-4 (5ng/ml) and anti-IFN gamma (1µg/ml) were used to direct to Th2 and IL-10 at 5ng/ml was used to direct to Tr1. After 4-5 days, the activated Th1, Th2 and Tr1 lymphocytes were washed once in DMEM and expanded for 4-7 days in DMEM 5% FCS (Hyclone), 100µM non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5}
25 5 M (Gibco), 10mM Hepes (Gibco) and IL-2 (1ng/ml). Following this, the activated Th1, Th2 and Tr1 lymphocytes were re-stimulated for 5 days with anti-CD28/OKT3 and cytokines as described above, but with the addition of anti-CD95L (1µg/ml) to prevent apoptosis. After 4-5 days, the Th1, Th2 and Tr1 lymphocytes were washed and then expanded again with IL-2 for 4-7 days. Activated Th1 and Th2 lymphocytes were
30 maintained in this way for a maximum of three cycles. RNA was prepared from primary and secondary Th1, Th2 and Tr1 after 6 and 24 hours following the second and third

activations with plate bound anti-CD3 and anti-CD28 mAbs and 4 days into the second and third expansion cultures in Interleukin 2.

The following leukocyte cells lines were obtained from the ATCC: Ramos, EOL-1, KU-812. EOL cells were further differentiated by culture in 0.1mM dbcAMP at
5 5×10^5 cells/ml for 8 days, changing the media every 3 days and adjusting the cell concentration to 5×10^5 cells/ml. For the culture of these cells, we used DMEM or RPMI (as recommended by the ATCC), with the addition of 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), 10mM Hepes (Gibco). RNA was either prepared from resting cells or cells
10 activated with PMA at 10ng/ml and ionomycin at 1 μ g/ml for 6 and 14 hours. Keratinocyte line CCD106 and an airway epithelial tumor line NCI-H292 were also obtained from the ATCC. Both were cultured in DMEM 5% FCS (Hyclone), 100 μ M non essential amino acids (Gibco), 1mM sodium pyruvate (Gibco), mercaptoethanol 5.5×10^{-5} M (Gibco), and 10mM Hepes (Gibco). CCD106 cells were activated for 6 and 14 hours with
15 approximately 5 ng/ml TNF alpha and 1ng/ml IL-1 beta, while NCI-H292 cells were activated for 6 and 14 hours with the following cytokines: 5ng/ml IL-4, 5ng/ml IL-9, 5ng/ml IL-13 and 25ng/ml IFN gamma.

For these cell lines and blood cells, RNA was prepared by lysing approximately 10^7 cells/ml using Trizol (Gibco BRL). Briefly, 1/10 volume of bromochloropropane
20 (Molecular Research Corporation) was added to the RNA sample, vortexed and after 10 minutes at room temperature, the tubes were spun at 14,000 rpm in a Sorvall SS34 rotor. The aqueous phase was removed and placed in a 15ml Falcon Tube. An equal volume of isopropanol was added and left at -20°C overnight. The precipitated RNA was spun down at 9,000 rpm for 15 min in a Sorvall SS34 rotor and washed in 70% ethanol. The pellet
25 was redissolved in 300 μ l of RNase-free water and 35 μ l buffer (Promega) 5 μ l DTT, 7 μ l RNasin and 8 μ l DNase were added. The tube was incubated at 37°C for 30 minutes to remove contaminating genomic DNA, extracted once with phenol chloroform and re-precipitated with 1/10 volume of 3M sodium acetate and 2 volumes of 100% ethanol. The RNA was spun down and placed in RNase free water. RNA was stored at -80°C.

30 AI_comprehensive panel_v1.0

The plates for AI_comprehensive panel_v1.0 include two control wells and 89 test samples comprised of cDNA isolated from surgical and postmortem human tissues

obtained from the Backus Hospital and Clinomics (Frederick, MD). Total RNA was extracted from tissue samples from the Backus Hospital in the Facility at CuraGen. Total RNA from other tissues was obtained from Clinomics.

5 Joint tissues including synovial fluid, synovium, bone and cartilage were obtained from patients undergoing total knee or hip replacement surgery at the Backus Hospital. Tissue samples were immediately snap frozen in liquid nitrogen to ensure that isolated RNA was of optimal quality and not degraded. Additional samples of osteoarthritis and rheumatoid arthritis joint tissues were obtained from Clinomics. Normal control tissues were supplied by Clinomics and were obtained during autopsy of trauma victims.

10 Surgical specimens of psoriatic tissues and adjacent matched tissues were provided as total RNA by Clinomics. Two male and two female patients were selected between the ages of 25 and 47. None of the patients were taking prescription drugs at the time samples were isolated.

Surgical specimens of diseased colon from patients with ulcerative colitis and
15 Crohns disease and adjacent matched tissues were obtained from Clinomics. Bowel tissue from three female and three male Crohn's patients between the ages of 41-69 were used. Two patients were not on prescription medication while the others were taking dexamethasone, phenobarbital, or tylenol. Ulcerative colitis tissue was from three male and four female patients. Four of the patients were taking lebid and two were on
20 phenobarbital.

Total RNA from post mortem lung tissue from trauma victims with no disease or with emphysema, asthma or COPD was purchased from Clinomics. Emphysema patients ranged in age from 40-70 and all were smokers, this age range was chosen to focus on patients with cigarette-linked emphysema and to avoid those patients with alpha-1-anti-
25 trypsin deficiencies. Asthma patients ranged in age from 36-75, and excluded smokers to prevent those patients that could also have COPD. COPD patients ranged in age from 35-80 and included both smokers and non-smokers. Most patients were taking corticosteroids, and bronchodilators.

In the labels employed to identify tissues in the AI_comprehensive panel_v1.0
30 panel, the following abbreviations are used:

AI = Autoimmunity

Syn = Synovial

Normal = No apparent disease

Rep22 /Rep20 = individual patients

RA = Rheumatoid arthritis

Backus = From Backus Hospital

5 OA = Osteoarthritis

(SS) (BA) (MF) = Individual patients

Adj = Adjacent tissue

Match control = adjacent tissues

-M = Male

10 -F = Female

COPD = Chronic obstructive pulmonary disease

Panels 5D and 5I

The plates for Panel 5D and 5I include two control wells and a variety of cDNAs isolated from human tissues and cell lines with an emphasis on metabolic diseases.

15 Metabolic tissues were obtained from patients enrolled in the Gestational Diabetes study. Cells were obtained during different stages in the differentiation of adipocytes from human mesenchymal stem cells. Human pancreatic islets were also obtained.

In the Gestational Diabetes study subjects are young (18 - 40 years), otherwise healthy women with and without gestational diabetes undergoing routine (elective)

20 Caesarean section. After delivery of the infant, when the surgical incisions were being repaired/closed, the obstetrician removed a small sample (<1 cc) of the exposed metabolic tissues during the closure of each surgical level. The biopsy material was rinsed in sterile saline, blotted and fast frozen within 5 minutes from the time of removal. The tissue was then flash frozen in liquid nitrogen and stored, individually, in sterile screw-top tubes and
25 kept on dry ice for shipment to or to be picked up by CuraGen. The metabolic tissues of interest include uterine wall (smooth muscle), visceral adipose, skeletal muscle (rectus) and subcutaneous adipose. Patient descriptions are as follows:

Patient 2: Diabetic Hispanic, overweight, not on insulin

Patient 7-9: Nondiabetic Caucasian and obese (BMI>30)

30 Patient 10: Diabetic Hispanic, overweight, on insulin

Patient 11: Nondiabetic African American and overweight

Patient 12: Diabetic Hispanic on insulin

Adiocyte differentiation was induced in donor progenitor cells obtained from Osirus (a division of Clonetics/BioWhittaker) in triplicate, except for Donor 3U which had only two replicates. Scientists at Clonetics isolated, grew and differentiated human
5 mesenchymal stem cells (HuMSCs) for CuraGen based on the published protocol found in Mark F. Pittenger, et al., Multilineage Potential of Adult Human Mesenchymal Stem Cells Science Apr 2 1999: 143-147. Clonetics provided Trizol lysates or frozen pellets suitable for mRNA isolation and ds cDNA production. A general description of each donor is as follows:

- 10 Donor 2 and 3 U: Mesenchymal Stem cells, Undifferentiated Adipose
Donor 2 and 3 AM: Adipose, AdiposeMidway Differentiated
Donor 2 and 3 AD: Adipose, Adipose Differentiated

Human cell lines were generally obtained from ATCC (American Type Culture
15 Collection), NCI or the German tumor cell bank and fall into the following tissue groups: kidney proximal convoluted tubule, uterine smooth muscle cells, small intestine, liver HepG2 cancer cells, heart primary stromal cells, and adrenal cortical adenoma cells. These cells are all cultured under standard recommended conditions and RNA extracted using the standard procedures. All samples were processed at CuraGen to produce single
20 stranded cDNA.

Panel 5I contains all samples previously described with the addition of pancreatic islets from a 58 year old female patient obtained from the Diabetes Research Institute at the University of Miami School of Medicine. Islet tissue was processed to total RNA at an outside source and delivered to CuraGen for addition to panel 5I.

25 In the labels employed to identify tissues in the 5D and 5I panels, the following abbreviations are used:

- GO Adipose = Greater Omentum Adipose
SK = Skeletal Muscle
UT = Uterus
30 PL = Placenta
AD = Adipose Differentiated
AM = Adipose Midway Differentiated

U = Undifferentiated Stem Cells

Panel CNSD.01

The plates for Panel CNSD.01 include two control wells and 94 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the
 5 Harvard Brain Tissue Resource Center. Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains two brains
 10 from each of the following diagnoses: Alzheimer's disease, Parkinson's disease, Huntington's disease, Progressive Supranuclear Palsy, Depression, and "Normal controls". Within each of these brains, the following regions are represented: cingulate gyrus, temporal pole, globus palladus, substantia nigra, Brodman Area 4 (primary motor strip), Brodman Area 7 (parietal cortex), Brodman Area 9 (prefrontal cortex), and Brodman area
 15 17 (occipital cortex). Not all brain regions are represented in all cases; e.g., Huntington's disease is characterized in part by neurodegeneration in the globus palladus, thus this region is impossible to obtain from confirmed Huntington's cases. Likewise Parkinson's disease is characterized by degeneration of the substantia nigra making this region more difficult to obtain. Normal control brains were examined for neuropathology and found to
 20 be free of any pathology consistent with neurodegeneration.

In the labels employed to identify tissues in the CNS panel, the following abbreviations are used:

PSP = Progressive supranuclear palsy

Sub Nigra = Substantia nigra

25 Glob Palladus= Globus palladus

Temp Pole = Temporal pole

Cing Gyr = Cingulate gyrus

BA 4 = Brodman Area 4

Panel CNS_Neurodegeneration_V1.0

30 The plates for Panel CNS_Neurodegeneration_V1.0 include two control wells and 47 test samples comprised of cDNA isolated from postmortem human brain tissue obtained from the Harvard Brain Tissue Resource Center (McLean Hospital) and the

Human Brain and Spinal Fluid Resource Center (VA Greater Los Angeles Healthcare System). Brains are removed from calvaria of donors between 4 and 24 hours after death, sectioned by neuroanatomists, and frozen at -80°C in liquid nitrogen vapor. All brains are sectioned and examined by neuropathologists to confirm diagnoses with clear associated neuropathology.

Disease diagnoses are taken from patient records. The panel contains six brains from Alzheimer's disease (AD) patients, and eight brains from "Normal controls" who showed no evidence of dementia prior to death. The eight normal control brains are divided into two categories: Controls with no dementia and no Alzheimer's like pathology (Controls) and controls with no dementia but evidence of severe Alzheimer's like pathology, (specifically senile plaque load rated as level 3 on a scale of 0-3; 0 = no evidence of plaques, 3 = severe AD senile plaque load). Within each of these brains, the following regions are represented: hippocampus, temporal cortex (Brodmann Area 21), parietal cortex (Brodmann area 7), and occipital cortex (Brodmann area 17). These regions were chosen to encompass all levels of neurodegeneration in AD. The hippocampus is a region of early and severe neuronal loss in AD; the temporal cortex is known to show neurodegeneration in AD after the hippocampus; the parietal cortex shows moderate neuronal death in the late stages of the disease; the occipital cortex is spared in AD and therefore acts as a "control" region within AD patients. Not all brain regions are represented in all cases.

In the labels employed to identify tissues in the CNS_Neurodegeneration_V1.0 panel, the following abbreviations are used:

AD = Alzheimer's disease brain; patient was demented and showed AD-like pathology upon autopsy

Control = Control brains; patient not demented, showing no neuropathology

Control (Path) = Control brains; patient not demented but showing severe AD-like pathology

SupTemporal Ctx = Superior Temporal Cortex

Inf Temporal Ctx = Inferior Temporal Cortex

A. CG100126-01: Keratin Associated Protein

Expression of gene CG100126-01 was assessed using the primer-probe set Ag4163, described in Table AA. Results of the RTQ-PCR runs are shown in Tables AB, AC and AD.

Table AA. Probe Name Ag4163

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - cagtgcgtgccagtctgtgt - 3'	19	217	203
Probe	TET-5' - agctgcagcatctccagctgctg - 3' - TAMRA	23	517	204
Reverse	5' - agctggattcacagcaagag - 3'	20	546	205

5

Table AB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4163, Run 221000456	Tissue Name	Rel. Exp.(%) Ag4163, Run 221000456
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	5.3
Melanoma* Hs688(B).T	2.4	Gastric ca. (liver met.) NCI-N87	3.4
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	100.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	8.1
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	95.9
Prostate Pool	0.0	Colon ca. CaCo-2	2.9
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	5.8	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	8.1
Ovarian ca. OVCAR-5	26.1	Small Intestine Pool	17.4
Ovarian ca. IGROV-1	31.2	Stomach Pool	3.5
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	4.2
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	9.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	81.8	Skeletal Muscle Pool	3.1

Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	2.6	Thymus Pool	1.6
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	5.7	CNS cancer (glio/astro) U-118-MG	10.2
Fetal Lung	0.0	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	8.2	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	31.9
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	26.6	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.0	Brain (whole)	0.0
Liver ca. HepG2	0.0	Spinal Cord Pool	0.0
Kidney Pool	9.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	3.2	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	5.6

Table AC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4163, Run 173333942	Tissue Name	Rel. Exp.(%) Ag4163, Run 173333942
Secondary Th1 act	0.0	HUVEC IL-1beta	0.1
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.3
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.1
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.1	Lung Microvascular EC	0.0

		TNFalpha + IL-1beta	
Primary Th2 act	0.0	Microvascular Dermal EC none	26.4
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.3
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL- 1beta	0.2
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1 _anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.3
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.1
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.1
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.2
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.4
PBMC PWM	0.0	Lung fibroblast IL-4	0.2
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.1
Ramos (B cell) none	0.0	Lung fibroblast IL-13	1.2
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0

B lymphocytes PWM	0.1	Dermal fibroblast CCD1070 rest	0.1
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.5
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.9
Dendritic cells anti-CD40	0.1	Neutrophils TNFa+LPS	0.7
Monocytes rest	0.0	Neutrophils rest	0.6
Monocytes LPS	0.0	Colon	0.8
Macrophages rest	0.1	Lung	3.6
Macrophages LPS	0.0	Thymus	15.5
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

Table AD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4163, Run 268624200	Tissue Name	Rel. Exp.(%) Ag4163, Run 268624200
Colon cancer 1	0.0	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	0.0
Colon cancer 2	0.0	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	0.6
Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	0.2
Colon malignant cancer 4	0.3	Adenocarcinoma of the prostate 4	0.5
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	0.3	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.1	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.7
Lung NAT 2	0.3	Adenocarcinoma of the prostate 9	0.0

Squamous cell carcinoma 3	0.2	Prostate cancer NAT 10	0.2
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	0.3	Kidney NAT 1	0.0
Melanoma 2	100.0	Kidney cancer 2	1.8
Melanoma 3	0.3	Kidney NAT 2	0.5
metastatic melanoma 4	0.2	Kidney cancer 3	0.0
metastatic melanoma 5	0.5	Kidney NAT 3	0.5
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4163 Results from one experiment with the CG100126-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4163 Highest expression of the CG100126-01 gene is detected in Melanoma LOXIMVI and a colon cancer HCT-116 cell lines (CTs=33.8). In addition, significant expression of this gene is also detected in a breast cancer T47D cell line. Therefore, expression of this gene can be used to distinguish these samples from other samples used in this panel and also as diagnostic marker for melanoma, colon and breast cancer. Furthermore, therapeutic modulation of this gene product can be useful in the treatment of these cancers.

Panel 4.1D Summary: Ag4163 Highest expression of the CG100126-01 gene is detected in kidney (CT=29). Moderate expression of this gene is also seen in lung, thymus and microvascular dermal endothelial cells. Therefore, expression of this gene can be used to distinguish these samples from other samples in this panel. In addition, therapeutic modulation of this gene can be useful in the treatment of autoimmune and inflammatory diseases that affect lung and kidney.

General oncology screening panel_v_2.4 Summary: Ag4163 Highest expression of the CG100126-01 gene is detected exclusively in melanoma sample(CT=29.8). Therefore, expression of this gene may be used as a diagnostic marker for detection of melanoma and therapeutic modulation of this gene product may be beneficial in the treatment of melanoma.

B. CG100146-01: UDP-Glucuronosyl Transferase

Expression of gene CG100146-01 was assessed using the primer-probe set Ag4165, described in Table BA. Results of the RTQ-PCR runs are shown in Tables BB and BC.

Table BA. Probe Name Ag4165

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agaccaggatgtctctgaaatg-3'	22	193	206
Probe	TET-5'-cgtcagtcctttctgctgatacagctca-3'-TAMRA	27	217	207
Reverse	5'-ccacaactcccagagctaaag-3'	21	251	208

5

Table BB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4165, Run 221000458	Tissue Name	Rel. Exp.(%) Ag4165, Run 221000458
Adipose	7.4	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	16.4
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	2.6
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	2.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	1.6	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	17.1
Placenta	0.0	Colon cancer tissue	4.4
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	1.9
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	100.0

Ovarian ca. IGROV-1	0.0	Stomach Pool	0.6
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	4.0	Fetal Heart	0.0
Breast ca. MCF-7	7.4	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	1.7
Breast Pool	0.8	Thymus Pool	2.8
Trachea	0.9	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.8	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	33.2	Brain (Thalamus) Pool	0.0
Fetal Liver	3.4	Brain (whole)	1.0
Liver ca. HepG2	0.0	Spinal Cord Pool	0.0
Kidney Pool	0.0	Adrenal Gland	1.4
Fetal Kidney	0.0	Pituitary gland Pool	0.0

Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	1.6

Table BC. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4165, Run 268624228	Tissue Name	Rel. Exp.(%) Ag4165, Run 268624228
Colon cancer 1	84.1	Bladder cancer NAT 2	0.0
Colon NAT 1	100.0	Bladder cancer NAT 3	0.0
Colon cancer 2	0.4	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	49.0	Adenocarcinoma of the prostate 1	0.0
Colon cancer 3	2.9	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	28.3	Adenocarcinoma of the prostate 3	0.0
Colon malignant cancer 4	19.3	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	59.9	Prostate cancer NAT 5	0.0
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	0.2	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	0.1	Kidney NAT 1	0.0
Melanoma 2	0.0	Kidney cancer 2	0.0
Melanoma 3	0.0	Kidney NAT 2	0.0
metastatic melanoma 4	0.0	Kidney cancer 3	0.1
metastatic melanoma 5	0.1	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer	0.0	Kidney NAT 4	0.0

NAT 1			
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4165 Expression of the CG100146-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4165 Highest expression of the CG100146-01 gene is detected exclusively in small intestine (Ct=32.7). In addition, low expression of this gene is also detected in liver. Thus, expression of this gene can be used to distinguish these samples from other samples used in this panel. Furthermore, therapeutic modulation of this gene product could be useful in the treatment of small intestine and liver related disorders. The CG100146-01 gene codes for UDP-glucuronosyltransferase. UDP-Glucuronosyltransferases (UGTs) are glycoproteins, which catalyze the conjugation of a broad variety of lipophilic aglycon substrates with glucuronic acid using UDP-glucuronic acid (UDP-GlcUA) as the sugar donor. The major function of glucuronidation is to change hydrophobic compounds into hydrophilic derivatives, a process which facilitates their detoxification and excretion (Radominska-Pandya et al., 2001, Curr Drug Metab 2(3):283-98, PMID: 11513331). Mutations in the UGT1A1 gene, one of the gene belonging to UGT family, are implicated in type I and type II Crigler-Najjar syndromes and the more common mild hyperbilirubinemia known as Gilbert syndrome (Kadakol et al., 2000, Hum. Mutat. 16: 297-306, PubMed ID:11013440). Thus, the CG100146-01 gene may also play a role in pathogenesis of Crigler-Najjar syndromes and Gilbert syndrome and therapeutic modulation of this gene could be beneficial in the treatment of these diseases.

Panel 4.1D Summary: Ag4165 Results from one experiment with the CG100146-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General oncology screening panel_v_2.4 Summary: Ag4165 Highest expression of the CG100146-01 gene is detected in colon sample (CT=27.8). Expression of this gene is exclusive to colon samples, with lower expression in cancer samples compared to adjacent normal tissue. Therefore, this gene may play a role as tumor suppressor and therapeutic modulation to increase the activity of the gene product may be beneficial in the treatment of colon cancer.

C. CG100179-01: Cyclophilin Like

Expression of gene CG100179-01 was assessed using the primer-probe set Ag4166, described in Table CA. Results of the RTQ-PCR runs are shown in Tables CB, CC, CD and CE.

5

Table CA. Probe Name Ag4166

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tttggacgagtgactaaaggaa-3'	22	2122	209
Probe	TET-5'-cagaggatctccaacgtcaaagtcaa-3'-TAMRA	26	2155	210
Reverse	5'-tcatagggcttatctgttttgg-3'	22	2183	211

Table CB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4166, Run 215538908	Tissue Name	Rel. Exp.(%) Ag4166, Run 215538908
AD 1 Hippo	12.2	Control (Path) 3 Temporal Ctx	7.6
AD 2 Hippo	24.3	Control (Path) 4 Temporal Ctx	52.5
AD 3 Hippo	10.2	AD 1 Occipital Ctx	27.2
AD 4 Hippo	10.7	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	97.3	AD 3 Occipital Ctx	14.2
AD 6 Hippo	61.1	AD 4 Occipital Ctx	21.8
Control 2 Hippo	18.9	AD 5 Occipital Ctx	37.6
Control 4 Hippo	11.3	AD 6 Occipital Ctx	28.9
Control (Path) 3 Hippo	11.3	Control 1 Occipital Ctx	8.2
AD 1 Temporal Ctx	23.0	Control 2 Occipital Ctx	41.2
AD 2 Temporal Ctx	29.3	Control 3 Occipital Ctx	30.4
AD 3 Temporal Ctx	11.7	Control 4 Occipital Ctx	9.7
AD 4 Temporal Ctx	26.8	Control (Path) 1 Occipital Ctx	76.3
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	18.7
AD 5 Sup Temporal Ctx	49.3	Control (Path) 3 Occipital Ctx	6.2
AD 6 Inf Temporal Ctx	68.8	Control (Path) 4 Occipital Ctx	25.9

AD 6 Sup Temporal Ctx	55.9	Control 1 Parietal Ctx	10.6
Control 1 Temporal Ctx	8.4	Control 2 Parietal Ctx	48.6
Control 2 Temporal Ctx	27.5	Control 3 Parietal Ctx	19.3
Control 3 Temporal Ctx	16.8	Control (Path) 1 Parietal Ctx	63.7
Control 3 Temporal Ctx	12.9	Control (Path) 2 Parietal Ctx	30.6
Control (Path) 1 Temporal Ctx	51.8	Control (Path) 3 Parietal Ctx	8.6
Control (Path) 2 Temporal Ctx	34.4	Control (Path) 4 Parietal Ctx	48.0

Table CC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4166, Run 221034286	Tissue Name	Rel. Exp.(%) Ag4166, Run 221034286
Adipose	9.7	Renal ca. TK-10	19.9
Melanoma* Hs688(A).T	9.3	Bladder	22.5
Melanoma* Hs688(B).T	8.0	Gastric ca. (liver met.) NCI-N87	44.8
Melanoma* M14	7.5	Gastric ca. KATO III	23.8
Melanoma* LOXIMVI	14.6	Colon ca. SW-948	5.8
Melanoma* SK-MEL-5	20.3	Colon ca. SW480	28.1
Squamous cell carcinoma SCC-4	7.5	Colon ca.* (SW480 met) SW620	29.7
Testis Pool	9.1	Colon ca. HT29	14.3
Prostate ca.* (bone met) PC-3	7.0	Colon ca. HCT-116	29.1
Prostate Pool	8.4	Colon ca. CaCo-2	30.1
Placenta	1.7	Colon cancer tissue	13.6
Uterus Pool	9.6	Colon ca. SW1116	5.0
Ovarian ca. OVCAR-3	16.4	Colon ca. Colo-205	5.3
Ovarian ca. SK-OV-3	28.9	Colon ca. SW-48	4.9
Ovarian ca. OVCAR-4	4.0	Colon Pool	20.6

Ovarian ca. OVCA-5	45.4	Small Intestine Pool	26.2
Ovarian ca. IGROV-1	11.1	Stomach Pool	14.8
Ovarian ca. OVCA-8	15.2	Bone Marrow Pool	9.9
Ovary	12.0	Fetal Heart	12.2
Breast ca. MCF-7	19.8	Heart Pool	8.5
Breast ca. MDA-MB-231	42.3	Lymph Node Pool	19.3
Breast ca. BT 549	40.9	Fetal Skeletal Muscle	8.2
Breast ca. T47D	100.0	Skeletal Muscle Pool	9.5
Breast ca. MDA-N	14.3	Spleen Pool	18.4
Breast Pool	18.0	Thymus Pool	22.4
Trachea	5.6	CNS cancer (glio/astro) U87-MG	36.1
Lung	19.5	CNS cancer (glio/astro) U-118-MG	27.9
Fetal Lung	26.2	CNS cancer (neuro;met) SK-N-AS	26.6
Lung ca. NCI-N417	3.9	CNS cancer (astro) SF-539	8.1
Lung ca. LX-1	19.8	CNS cancer (astro) SNB-75	49.0
Lung ca. NCI-H146	5.1	CNS cancer (glio) SNB-19	10.9
Lung ca. SHP-77	17.3	CNS cancer (glio) SF-295	46.0
Lung ca. A549	20.0	Brain (Amygdala) Pool	8.6
Lung ca. NCI-H526	3.3	Brain (cerebellum)	9.7
Lung ca. NCI-H23	20.9	Brain (fetal)	12.9
Lung ca. NCI-H460	15.7	Brain (Hippocampus) Pool	6.8
Lung ca. HOP-62	5.0	Cerebral Cortex Pool	13.5
Lung ca. NCI-H522	26.8	Brain (Substantia nigra) Pool	7.9
Liver	0.7	Brain (Thalamus) Pool	14.0
Fetal Liver	14.5	Brain (whole)	4.7
Liver ca. HepG2	11.4	Spinal Cord Pool	10.2
Kidney Pool	45.1	Adrenal Gland	3.8

Fetal Kidney	27.5	Pituitary gland Pool	3.6
Renal ca. 786-0	9.5	Salivary Gland	0.8
Renal ca. A498	6.0	Thyroid (female)	6.2
Renal ca. ACHN	7.3	Pancreatic ca. CAPAN2	25.2
Renal ca. UO-31	11.7	Pancreas Pool	21.0

Table CD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4166, Run 173334537	Tissue Name	Rel. Exp.(%) Ag4166, Run 173334537
Secondary Th1 act	77.9	HUVEC IL-1beta	49.0
Secondary Th2 act	74.2	HUVEC IFN gamma	45.4
Secondary Tr1 act	66.0	HUVEC TNF alpha + IFN gamma	29.3
Secondary Th1 rest	39.8	HUVEC TNF alpha + IL4	33.4
Secondary Th2 rest	55.5	HUVEC IL-11	32.5
Secondary Tr1 rest	58.2	Lung Microvascular EC none	57.8
Primary Th1 act	68.8	Lung Microvascular EC TNFalpha + IL-1beta	39.8
Primary Th2 act	94.0	Microvascular Dermal EC none	33.7
Primary Tr1 act	81.2	Microvascular Dermal EC TNFalpha + IL-1beta	23.0
Primary Th1 rest	55.5	Bronchial epithelium TNFalpha + IL1beta	20.9
Primary Th2 rest	60.3	Small airway epithelium none	7.7
Primary Tr1 rest	66.0	Small airway epithelium TNFalpha + IL-1beta	18.7
CD45RA CD4 lymphocyte act	48.6	Coronary artery SMC rest	17.8
CD45RO CD4 lymphocyte act	87.1	Coronary artery SMC TNFalpha + IL-1beta	22.1
CD8 lymphocyte act	66.4	Astrocytes rest	14.1
Secondary CD8 lymphocyte rest	62.9	Astrocytes TNFalpha + IL-1beta	12.5
Secondary CD8 lymphocyte act	33.4	KU-812 (Basophil) rest	37.6
CD4 lymphocyte	63.3	KU-812 (Basophil)	45.7

none		PMA/ionomycin	
2ry Th1/Th2/Tr1_anti- CD95 CH11	100.0	CCD1106 (Keratinocytes) none	43.5
LAK cells rest	33.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	27.2
LAK cells IL-2	62.4	Liver cirrhosis	20.4
LAK cells IL- 2+IL-12	61.1	NCI-H292 none	29.5
LAK cells IL- 2+IFN gamma	55.9	NCI-H292 IL-4	34.4
LAK cells IL-2+ IL-18	67.4	NCI-H292 IL-9	54.7
LAK cells PMA/ionomycin	18.2	NCI-H292 IL-13	59.9
NK Cells IL-2 rest	91.4	NCI-H292 IFN gamma	33.9
Two Way MLR 3 day	58.2	HPAEC none	27.5
Two Way MLR 5 day	52.5	HPAEC TNF alpha + IL-1 beta	34.4
Two Way MLR 7 day	35.4	Lung fibroblast none	35.4
PBMC rest	17.0	Lung fibroblast TNF alpha + IL-1 beta	29.9
PBMC PWM	53.6	Lung fibroblast IL-4	28.9
PBMC PHA-L	50.0	Lung fibroblast IL-9	25.9
Ramos (B cell) none	61.1	Lung fibroblast IL-13	24.1
Ramos (B cell) ionomycin	39.8	Lung fibroblast IFN gamma	24.1
B lymphocytes PWM	35.4	Dermal fibroblast CCD1070 rest	49.7
B lymphocytes CD40L and IL-4	71.2	Dermal fibroblast CCD1070 TNF alpha	71.7
EOL-1 dbcAMP	60.3	Dermal fibroblast CCD1070 IL-1 beta	35.8
EOL-1 dbcAMP PMA/ionomycin	60.3	Dermal fibroblast IFN gamma	29.5
Dendritic cells none	26.6	Dermal fibroblast IL-4	40.1
Dendritic cells LPS	20.0	Dermal Fibroblasts rest	27.9
Dendritic cells	22.5	Neutrophils TNFa+LPS	5.3

anti-CD40			
Monocytes rest	33.7	Neutrophils rest	33.7
Monocytes LPS	35.4	Colon	16.4
Macrophages rest	31.2	Lung	14.9
Macrophages LPS	14.2	Thymus	62.0
HUVEC none	27.7	Kidney	60.3
HUVEC starved	59.0		

Table CE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4166, Run 268624305	Tissue Name	Rel. Exp.(%) Ag4166, Run 268624305
Colon cancer 1	20.0	Bladder cancer NAT 2	3.5
Colon cancer NAT 1	7.5	Bladder cancer NAT 3	0.3
Colon cancer 2	12.9	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	5.3	Adenocarcinoma of the prostate 1	32.1
Colon cancer 3	34.4	Adenocarcinoma of the prostate 2	3.5
Colon cancer NAT 3	17.4	Adenocarcinoma of the prostate 3	8.1
Colon malignant cancer 4	39.2	Adenocarcinoma of the prostate 4	19.2
Colon normal adjacent tissue 4	2.7	Prostate cancer NAT 5	9.2
Lung cancer 1	15.5	Adenocarcinoma of the prostate 6	2.4
Lung NAT 1	3.9	Adenocarcinoma of the prostate 7	4.5
Lung cancer 2	55.5	Adenocarcinoma of the prostate 8	1.3
Lung NAT 2	5.8	Adenocarcinoma of the prostate 9	16.4
Squamous cell carcinoma 3	13.7	Prostate cancer NAT 10	0.9
Lung NAT 3	0.5	Kidney cancer 1	33.0
metastatic melanoma 1	23.2	Kidney NAT 1	12.5
Melanoma 2	2.9	Kidney cancer 2	100.0
Melanoma 3	8.2	Kidney NAT 2	26.4
metastatic melanoma 4	43.5	Kidney cancer 3	40.9

metastatic melanoma 5	66.0	Kidney NAT 3	11.3
Bladder cancer 1	7.9	Kidney cancer 4	10.9
Bladder cancer NAT 1	0.0	Kidney NAT 4	4.0
Bladder cancer 2	5.0		

CNS_neurodegeneration_v1.0 Summary: Ag4166 This panel confirms the expression of the CG100179-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4166 Highest expression of the CG100179-01 gene is detected in the breast cancer T47D cell line (CT=26). High expression of this gene is seen in cluster of breast, ovarian, pancreatic, CNS, colon, gastric, renal, lung cancer cell lines and melanoma cell lines. Thus, therapeutic modulation of this gene could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4166 Highest expression of the CG100179-01 gene is detected in anti-CD95 CH11 treated secondary Th1/Th2/Tr1 cells (CT=28.5). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression

suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4166 Highest expression of the CG100179-01 gene is detected in kidney cancer (CT=28.5), with significant expression also seen in melanoma, colon, lung, bladder, prostate and kidney cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene could be as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

D. CG100212-01 and CG100212-02: Novel Gene Belonging to Zinc-Containing Alcohol Dehydrogenase Superfamily

Expression of gene CG100212-01 and full length physical clone CG100212-02 was assessed using the primer-probe set Ag4167, described in Table DA. Results of the RTQ-PCR runs are shown in Tables DB, DC, DD and DE. Please note that CG100212-02 represents a full-length physical clone of the CG100212-01 gene, validating the prediction of the gene sequence.

Table DA. Probe Name Ag4167

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggatccttacatgcgtttaga-3'	22	146	212
Probe	TET-5'-tggcactgattatataacaccttggca-3'-TAMRA	27	182	213
Reverse	5'-tccatcaacgacttgagatagc-3'	22	209	214

Table DB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4167, Run 215538909	Tissue Name	Rel. Exp.(%) Ag4167, Run 215538909
AD 1 Hippo	18.0	Control (Path) 3 Temporal	9.9

		Ctx	
AD 2 Hippo	44.4	Control (Path) 4 Temporal Ctx	47.3
AD 3 Hippo	10.7	AD 1 Occipital Ctx	21.9
AD 4 Hippo	12.9	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	56.3	AD 3 Occipital Ctx	9.3
AD 6 Hippo	67.8	AD 4 Occipital Ctx	24.5
Control 2 Hippo	31.0	AD 5 Occipital Ctx	15.9
Control 4 Hippo	25.3	AD 6 Occipital Ctx	34.6
Control (Path) 3 Hippo	13.3	Control 1 Occipital Ctx	9.2
AD 1 Temporal Ctx	34.2	Control 2 Occipital Ctx	39.8
AD 2 Temporal Ctx	100.0	Control 3 Occipital Ctx	22.2
AD 3 Temporal Ctx	9.4	Control 4 Occipital Ctx	13.8
AD 4 Temporal Ctx	15.1	Control (Path) 1 Occipital Ctx	87.7
AD 5 Inf Temporal Ctx	79.6	Control (Path) 2 Occipital Ctx	17.6
AD 5 Sup Temporal Ctx	51.1	Control (Path) 3 Occipital Ctx	5.6
AD 6 Inf Temporal Ctx	71.7	Control (Path) 4 Occipital Ctx	16.8
AD 6 Sup Temporal Ctx	60.3	Control 1 Parietal Ctx	18.4
Control 1 Temporal Ctx	10.2	Control 2 Parietal Ctx	48.6
Control 2 Temporal Ctx	34.9	Control 3 Parietal Ctx	14.0
Control 3 Temporal Ctx	29.5	Control (Path) 1 Parietal Ctx	75.3
Control 4 Temporal Ctx	16.6	Control (Path) 2 Parietal Ctx	26.8
Control (Path) 1 Temporal Ctx	61.6	Control (Path) 3 Parietal Ctx	10.2
Control (Path) 2 Temporal Ctx	36.9	Control (Path) 4 Parietal Ctx	42.0

Table DC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4167, Run 221035838	Tissue Name	Rel. Exp.(%) Ag4167, Run
-------------	------------------------------------	-------------	--------------------------

			221035838
Adipose	3.6	Renal ca. TK-10	22.5
Melanoma* Hs688(A).T	5.3	Bladder	5.4
Melanoma* Hs688(B).T	5.8	Gastric ca. (liver met.) NCI-N87	18.6
Melanoma* M14	1.6	Gastric ca. KATO III	31.9
Melanoma* LOXIMVI	5.0	Colon ca. SW-948	7.3
Melanoma* SK-MEL-5	12.5	Colon ca. SW480	22.7
Squamous cell carcinoma SCC-4	17.3	Colon ca.* (SW480 met) SW620	14.0
Testis Pool	7.0	Colon ca. HT29	4.7
Prostate ca.* (bone met) PC-3	27.0	Colon ca. HCT-116	25.7
Prostate Pool	6.8	Colon ca. CaCo-2	24.1
Placenta	0.7	Colon cancer tissue	7.4
Uterus Pool	2.7	Colon ca. SW1116	2.8
Ovarian ca. OVCAR-3	28.3	Colon ca. Colo-205	3.3
Ovarian ca. SK-OV-3	30.4	Colon ca. SW-48	1.7
Ovarian ca. OVCAR-4	9.5	Colon Pool	5.8
Ovarian ca. OVCAR-5	59.5	Small Intestine Pool	8.8
Ovarian ca. IGROV-1	8.7	Stomach Pool	5.3
Ovarian ca. OVCAR-8	3.9	Bone Marrow Pool	2.8
Ovary	4.4	Fetal Heart	13.7
Breast ca. MCF-7	27.2	Heart Pool	8.5
Breast ca. MDA-MB-231	17.7	Lymph Node Pool	6.7
Breast ca. BT 549	54.7	Fetal Skeletal Muscle	5.2
Breast ca. T47D	100.0	Skeletal Muscle Pool	22.2
Breast ca. MDA-N	3.3	Spleen Pool	4.2
Breast Pool	6.9	Thymus Pool	4.5
Trachea	7.6	CNS cancer (glio/astro) U87-MG	24.8
Lung	3.7	CNS cancer (glio/astro) U-	17.4

		118-MG	
Fetal Lung	10.7	CNS cancer (neuro;met) SK-N-AS	9.2
Lung ca. NCI-N417	4.1	CNS cancer (astro) SF-539	5.6
Lung ca. LX-1	27.4	CNS cancer (astro) SNB-75	16.0
Lung ca. NCI-H146	1.2	CNS cancer (glio) SNB-19	7.9
Lung ca. SHP-77	19.5	CNS cancer (glio) SF-295	15.2
Lung ca. A549	20.6	Brain (Amygdala) Pool	6.7
Lung ca. NCI-H526	1.0	Brain (cerebellum)	6.3
Lung ca. NCI-H23	33.9	Brain (fetal)	7.8
Lung ca. NCI-H460	8.0	Brain (Hippocampus) Pool	8.3
Lung ca. HOP-62	8.5	Cerebral Cortex Pool	8.1
Lung ca. NCI-H522	27.7	Brain (Substantia nigra) Pool	7.0
Liver	1.3	Brain (Thalamus) Pool	12.8
Fetal Liver	8.4	Brain (whole)	6.8
Liver ca. HepG2	7.9	Spinal Cord Pool	10.6
Kidney Pool	10.0	Adrenal Gland	10.5
Fetal Kidney	11.8	Pituitary gland Pool	2.8
Renal ca. 786-0	9.8	Salivary Gland	3.0
Renal ca. A498	2.8	Thyroid (female)	5.4
Renal ca. ACHN	9.4	Pancreatic ca. CAPAN2	24.3
Renal ca. UO-31	15.7	Pancreas Pool	8.0

Table DD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4167, Run 173350762	Tissue Name	Rel. Exp.(%) Ag4167, Run 173350762
Secondary Th1 act	8.5	HUVEC IL-1beta	9.4
Secondary Th2 act	10.7	HUVEC IFN gamma	14.0
Secondary Tr1 act	5.3	HUVEC TNF alpha + IFN gamma	4.3
Secondary Th1 rest	6.2	HUVEC TNF alpha + IL4	6.1
Secondary Th2 rest	11.5	HUVEC IL-11	6.9
Secondary Tr1 rest	7.4	Lung Microvascular EC none	15.3

Primary Th1 act	14.0	Lung Microvascular EC TNFalpha + IL-1beta	10.2
Primary Th2 act	9.3	Microvascular Dermal EC none	17.1
Primary Tr1 act	9.2	Microvascular Dermal EC TNFalpha + IL-1beta	6.3
Primary Th1 rest	4.4	Bronchial epithelium TNFalpha + IL1beta	28.1
Primary Th2 rest	3.0	Small airway epithelium none	9.9
Primary Tr1 rest	5.9	Small airway epithelium TNFalpha + IL-1beta	31.4
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	20.4
CD45RO CD4 lymphocyte act	22.8	Coronary artery SMC TNFalpha + IL-1beta	14.1
CD8 lymphocyte act	16.8	Astrocytes rest	21.2
Secondary CD8 lymphocyte rest	12.4	Astrocytes TNFalpha + IL- 1beta	8.6
Secondary CD8 lymphocyte act	12.3	KU-812 (Basophil) rest	22.1
CD4 lymphocyte none	4.7	KU-812 (Basophil) PMA/ionomycin	31.4
2ry Th1/Th2/Tr1_anti- CD95 CH11	3.4	CCD1106 (Keratinocytes) none	38.4
LAK cells rest	21.2	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	30.6
LAK cells IL-2	7.0	Liver cirrhosis	23.7
LAK cells IL- 2+IL-12	6.7	NCI-H292 none	66.0
LAK cells IL- 2+IFN gamma	4.6	NCI-H292 IL-4	61.6
LAK cells IL-2+ IL-18	11.7	NCI-H292 IL-9	100.0
LAK cells PMA/ionomycin	17.7	NCI-H292 IL-13	59.5
NK Cells IL-2 rest	9.0	NCI-H292 IFN gamma	68.3
Two Way MLR 3 day	14.8	HPAEC none	12.2
Two Way MLR 5 day	14.5	HPAEC TNF alpha + IL-1 beta	23.0
Two Way MLR 7	13.6	Lung fibroblast none	21.5

day			
PBMC rest	4.5	Lung fibroblast TNF alpha + IL-1 beta	29.9
PBMC PWM	13.2	Lung fibroblast IL-4	15.2
PBMC PHA-L	20.6	Lung fibroblast IL-9	19.2
Ramos (B cell) none	0.7	Lung fibroblast IL-13	21.9
Ramos (B cell) ionomycin	1.1	Lung fibroblast IFN gamma	13.1
B lymphocytes PWM	22.5	Dermal fibroblast CCD1070 rest	9.8
B lymphocytes CD40L and IL-4	6.7	Dermal fibroblast CCD1070 TNF alpha	12.8
EOL-1 dbcAMP	5.5	Dermal fibroblast CCD1070 IL-1 beta	7.0
EOL-1 dbcAMP PMA/ionomycin	1.4	Dermal fibroblast IFN gamma	8.4
Dendritic cells none	26.8	Dermal fibroblast IL-4	28.1
Dendritic cells LPS	27.2	Dermal Fibroblasts rest	16.3
Dendritic cells anti-CD40	25.7	Neutrophils TNFa+LPS	1.8
Monocytes rest	7.9	Neutrophils rest	1.0
Monocytes LPS	8.4	Colon	12.0
Macrophages rest	52.5	Lung	10.4
Macrophages LPS	14.6	Thymus	22.2
HUVEC none	6.5	Kidney	77.9
HUVEC starved	11.0		

Table DE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4167, Run 268624306	Tissue Name	Rel. Exp.(%) Ag4167, Run 268624306
Colon cancer 1	10.4	Bladder cancer NAT 2	0.7
Colon NAT 1	4.7	Bladder cancer NAT 3	0.2
Colon cancer 2	18.3	Bladder cancer NAT 4	2.7
Colon cancer NAT 2	10.5	Adenocarcinoma of the prostate 1	21.3
Colon cancer 3	22.8	Adenocarcinoma of the prostate 2	2.8
Colon cancer NAT 3	15.3	Adenocarcinoma of the prostate 3	19.2

Colon malignant cancer 4	28.1	Adenocarcinoma of the prostate 4	19.6
Colon normal adjacent tissue 4	5.6	Prostate cancer NAT 5	1.5
Lung cancer 1	13.0	Adenocarcinoma of the prostate 6	7.5
Lung NAT 1	1.0	Adenocarcinoma of the prostate 7	7.0
Lung cancer 2	41.2	Adenocarcinoma of the prostate 8	3.5
Lung NAT 2	2.9	Adenocarcinoma of the prostate 9	31.6
Squamous cell carcinoma 3	20.0	Prostate cancer NAT 10	3.1
Lung NAT 3	1.3	Kidney cancer 1	21.8
metastatic melanoma 1	16.8	Kidney NAT 1	11.4
Melanoma 2	1.3	Kidney cancer 2	100.0
Melanoma 3	4.8	Kidney NAT 2	43.5
metastatic melanoma 4	24.7	Kidney cancer 3	44.4
metastatic melanoma 5	35.8	Kidney NAT 3	14.3
Bladder cancer 1	2.9	Kidney cancer 4	17.6
Bladder cancer NAT 1	0.0	Kidney NAT 4	9.9
Bladder cancer 2	3.7		

CNS_neurodegeneration_v1.0 Summary: Ag4167 This panel confirms the expression of the CG100212-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4167 Highest expression of the CG100212-01 gene is detected in a breast cancer T47D cell line (CT=26). High expression of this gene is seen in cluster of breast, ovarian, colon, gastric, renal, lung, pancreatic, CNS, hepatic, prostate cancer cell lines and melanoma cell lines. Thus, therapeutic modulation of this gene product could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4167 Highest expression of the CG100212-01 gene is detected in IL-9 treated NCI-H292 cell line (CT=31). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4167 Highest expression of the CG100212-01 gene is detected in kidney cancer (CT=28.9), with significant expression also seen in metastatic melanoma, colon, lung, bladder, prostate and kidney cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene could be as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

E. CG100222-01: NADP-Dependent Leukotriene B4

Expression of gene CG100222-01 was assessed using the primer-probe set Ag4169, described in Table EA. Results of the RTQ-PCR runs are shown in Tables EB and EC.

5

Table EA. Probe Name Ag4169

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gagcttcacatggaagggtt-3'	20	783	215
Probe	TET-5'-attgtgacctgctggccaggagat-3'-TAMRA	24	804	216
Reverse	5'-ctctgagacccatttcagca-3'	20	853	217

Table EB. General screening panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4169, Run 221035855	Tissue Name	Rel. Exp.(%) Ag4169, Run 221035855
Adipose	0.1	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.7
Melanoma* Hs688(B).T	0.4	Gastric ca. (liver met.) NCI-N87	2.7
Melanoma* M14	0.0	Gastric ca. KATO III	0.2
Melanoma* LOXIMVI	0.2	Colon ca. SW-948	0.0
Melanoma* SK- MEL-5	1.4	Colon ca. SW480	1.1
Squamous cell carcinoma SCC-4	0.3	Colon ca.* (SW480 met) SW620	0.6
Testis Pool	2.4	Colon ca. HT29	0.6
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.2
Prostate Pool	0.0	Colon ca. CaCo-2	4.6
Placenta	0.1	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.2
Ovarian ca. OVCAR-3	1.6	Colon ca. Colo-205	0.0
Ovarian ca. SK- OV-3	100.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	0.8	Small Intestine Pool	0.2

Ovarian ca. IGROV-1	0.4	Stomach Pool	0.3
Ovarian ca. OVCAR-8	0.2	Bone Marrow Pool	0.2
Ovary	0.0	Fetal Heart	0.1
Breast ca. MCF-7	0.3	Heart Pool	0.3
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.4
Breast ca. BT 549	0.3	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.7	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.2	Spleen Pool	0.0
Breast Pool	0.1	Thymus Pool	0.3
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.1	CNS cancer (glio/astro) U-118-MG	0.3
Fetal Lung	0.3	CNS cancer (neuro;met) SK-N-AS	2.3
Lung ca. NCI-N417	0.2	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.5	CNS cancer (astro) SNB-75	0.4
Lung ca. NCI-H146	0.2	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	1.6	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.4	Brain (cerebellum)	0.0
Lung ca. NCI-H23	3.7	Brain (fetal)	0.3
Lung ca. NCI-H460	0.5	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.2	Cerebral Cortex Pool	0.3
Lung ca. NCI-H522	0.4	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.1	Brain (whole)	0.2
Liver ca. HepG2	0.6	Spinal Cord Pool	0.0
Kidney Pool	0.6	Adrenal Gland	0.0
Fetal Kidney	1.1	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.2	Thyroid (female)	0.2
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.7
Renal ca. UO-31	0.0	Pancreas Pool	0.0

Table EC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4169, Run	Tissue Name	Rel. Exp.(%) Ag4169, Run 173333241
-------------	--------------------------------	-------------	--

	173333241		
Secondary Th1 act	0.0	HUVEC IL-1beta	0.4
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.4
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.4	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.5	Small airway epithelium none	0.0
Primary Tr1 rest	0.3	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.5	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.7	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.4	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.8	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.3
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.4
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	1.2
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	1.3
LAK cells IL-2+ IL-18	0.7	NCI-H292 IL-9	4.1
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.8
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.8
Two Way MLR 3 day	0.0	HPAEC none	0.3

Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.4	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.4
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.4
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.8
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.4	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	1.5
Dendritic cells none	0.4	Dermal fibroblast IL-4	0.9
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	1.5
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	1.3
Monocytes rest	0.0	Neutrophils rest	0.4
Monocytes LPS	0.0	Colon	1.3
Macrophages rest	0.0	Lung	3.0
Macrophages LPS	0.0	Thymus	7.6
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4169 Expression of the CG100222-01 gene is low/undetectable (CTs > 34.5) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4169 Highest expression of the CG100222-01 gene is detected in ovarian cancer SK-OV-3 cell line (CT=28). In addition, low but significant expression of this gene is seen in CNS cancer, colon cancer, lung cancer, ovarian cancer and melanoma cell lines. Thus, expression of this gene could be used as diagnostic marker in detection of these cancers. Therapeutic modulation of this gene product through the use of antibodies or small molecule drugs, may be beneficial in the treatment of these cancers. [mpattu, 25-Mar-02]

Panel 4.1D Summary: Ag4169 Highest expression of the CG100222-01 gene is detected in Kidney (CT=29). Thus, expression of this gene can be used to distinguish the

kidney sample from other samples used in this panel. In addition, low but significant expression of this gene is also seen in lung and thymus. Therefore, therapeutic modulation of this gene can be beneficial in the autoimmune and inflammatory diseases that affect lung and kidney.

- 5 **Panel 5 Islet Summary:** Ag4169 Expression of the CG100222-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4169 Results from one experiment with the CG100222-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

10

F. CG100266-01 and CG100266-02: Cyclophilin A Like Gene

Expression of gene CG100266-01 and full length clone CG100266-02 was assessed using the primer-probe set Ag4174, described in Table FA. Results of the RTQ-PCR runs are shown in Tables FB, FC and FD.

15

Table FA. Probe Name Ag4174

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gggagaaatttgatgacaagaa-3'	22	371	218
Probe	TET-5'-cttcatacctaaagcacgcaggtcctg-3'-TAMRA	26	393	219
Reverse	5'-actgtttgtgttgggtccagta-3'	22	438	220

Table FB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4174, Run 221034287	Tissue Name	Rel. Exp.(%) Ag4174, Run 221034287
Adipose	1.6	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	3.2
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	2.4	Colon ca. SW-948	0.0
Melanoma* SK- MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca. * (SW480 met) SW620	0.0
Testis Pool	14.0	Colon ca. HT29	0.0

Prostate ca.* (bone met) PC-3	2.9	Colon ca. HCT-116	5.3
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	7.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	2.3	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	2.3
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	1.6
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	3.7
Breast ca. T47D	5.5	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	3.5	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	0.0
Trachea	28.3	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	5.5
Fetal Lung	62.0	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	3.3
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	100.0	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	5.2
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	4.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	14.8

Lung ca. HOP-62	2.5	Cerebral Cortex Pool	1.8
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	1.2
Fetal Liver	0.0	Brain (whole)	4.4
Liver ca. HepG2	2.2	Spinal Cord Pool	3.6
Kidney Pool	0.9	Adrenal Gland	0.0
Fetal Kidney	18.6	Pituitary gland Pool	2.3
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	2.7

Table FC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4174, Run 173507624	Tissue Name	Rel. Exp.(%) Ag4174, Run 173507624
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.5
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8	0.0	KU-812 (Basophil) rest	0.0

lymphocyte act			
CD4 lymphocyte none	0.6	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	1.0
LAK cells IL-2+IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.6
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.6	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.5	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.2
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.4	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	1.4
Dendritic cells LPS	0.3	Dermal Fibroblasts rest	2.7

Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	1.0
Monocytes LPS	0.0	Colon	1.8
Macrophages rest	0.5	Lung	4.0
Macrophages LPS	0.0	Thymus	21.9
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

Table FD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4174, Run 268624376	Tissue Name	Rel. Exp.(%) Ag4174, Run 268624376
Colon cancer 1	0.0	Bladder cancer NAT 2	4.3
Colon cancer NAT 1	0.0	Bladder cancer NAT 3	0.0
Colon cancer 2	0.0	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	4.2
Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	5.6
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	4.9	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	14.2	Adenocarcinoma of the prostate 7	4.0
Lung cancer 2	5.2	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	8.8	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	5.2	Prostate cancer NAT 10	0.0
Lung NAT 3	9.7	Kidney cancer 1	0.0
metastatic melanoma 1	100.0	Kidney NAT 1	0.0
Melanoma 2	4.2	Kidney cancer 2	0.0
Melanoma 3	0.0	Kidney NAT 2	0.0
metastatic melanoma 4	0.0	Kidney cancer 3	0.0
metastatic melanoma 5	18.6	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	0.0

Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4174 Results from one experiment with the CG100266-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4174 Highest expression of the CG100266-01 gene is seen in a lung cancer cell line (CT=30.2). Significant levels of expression are also seen in fetal lung tissue, which is more proliferative than the adult counterpart. Thus, expression of this gene, a cyclophilin A homog. could be used to differentiate between fetal lung and adult lung and as a marker of lung cancer. This is in agreement with published reports that a novel cyclophilin like molecule is co-expressed in small cell lung cancer cell lines (Kim JO. Oncogene 1998 Aug 27;17(8):1019-26). Furthermore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of lung cancer.

Low but significant levels of expression are also seen in the amygdala, hippocampus and whole and fetal brain. Cyclophilin A has been implicated in neuronal differentiation and human embryonic brain cells (Nahreini P, Cell Mol Neurobiol 2001 Feb;21(1):65-79). Based on the expression of this cyclophilin homolog in the CNS, modulation of the expression of this gene may be useful in the treatment of neurodegenerative disorders, including Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 4.1D Summary: Ag4174 Expression of the CG100266-01 gene is restricted to normal kidney, lung and thymus (CTs=30-34). This pattern of expression suggests that this gene product may be involved in the normal homeostasis of these tissues. Therapeutic modulation of the expression or function of this gene may be useful in maintaining or restoring function to these organs during inflammation.

General oncology screening panel_v_2.4 Summary: Ag4174 Expression of the CG100266-01 gene is restricted to a melanoma (CT=33.8). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of melanoma. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of melanoma.

G. CG100456-01: CoA Transferase-Like

Expression of gene CG100456-01 was assessed using the primer-probe set Ag4179, described in Table GA. Results of the RTQ-PCR runs are shown in Tables GB, GC and GD.

5

Table GA. Probe Name Ag4179

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - ccaggtgtggaaatgaagaag - 3'	21	1030	221
Probe	TET- 5' - atcactgacacccacacgccgtct - 3' - TAMRA	24	1051	222
Reverse	5' - gcatgtaatagagcgtcagggtt - 3'	22	1087	223

Table GB. General screening panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4179, Run 221117873	Tissue Name	Rel. Exp.(%) Ag4179, Run 221117873
Adipose	2.4	Renal ca. TK-10	43.2
Melanoma* Hs688(A).T	18.8	Bladder	8.4
Melanoma* Hs688(B).T	20.2	Gastric ca. (liver met.) NCI-N87	23.5
Melanoma* M14	23.0	Gastric ca. KATO III	55.5
Melanoma* LOXIMVI	8.9	Colon ca. SW-948	19.8
Melanoma* SK- MEL-5	22.7	Colon ca. SW480	28.5
Squamous cell carcinoma SCC-4	22.1	Colon ca.* (SW480 met) SW620	15.3
Testis Pool	9.2	Colon ca. HT29	18.0
Prostate ca.* (bone met) PC-3	24.0	Colon ca. HCT-116	38.2
Prostate Pool	3.3	Colon ca. CaCo-2	24.1
Placenta	15.6	Colon cancer tissue	13.1
Uterus Pool	1.6	Colon ca. SW1116	12.0
Ovarian ca. OVCAR-3	28.7	Colon ca. Colo-205	16.4
Ovarian ca. SK- OV-3	36.3	Colon ca. SW-48	12.0
Ovarian ca. OVCAR-4	20.6	Colon Pool	8.0
Ovarian ca. OVCAR-5	20.3	Small Intestine Pool	6.0

Ovarian ca. IGROV-1	17.3	Stomach Pool	4.1
Ovarian ca. OVCAR-8	12.7	Bone Marrow Pool	2.6
Ovary	6.5	Fetal Heart	5.6
Breast ca. MCF-7	27.7	Heart Pool	4.1
Breast ca. MDA-MB-231	39.5	Lymph Node Pool	9.2
Breast ca. BT 549	100.0	Fetal Skeletal Muscle	4.8
Breast ca. T47D	36.6	Skeletal Muscle Pool	9.5
Breast ca. MDA-N	16.6	Spleen Pool	5.1
Breast Pool	7.5	Thymus Pool	6.6
Trachea	10.6	CNS cancer (glio/astro) U87-MG	16.0
Lung	0.8	CNS cancer (glio/astro) U-118-MG	34.6
Fetal Lung	9.6	CNS cancer (neuro;met) SK-N-AS	24.7
Lung ca. NCI-N417	15.2	CNS cancer (astro) SF-539	28.3
Lung ca. LX-1	21.0	CNS cancer (astro) SNB-75	53.6
Lung ca. NCI-H146	7.5	CNS cancer (glio) SNB-19	13.6
Lung ca. SHP-77	17.8	CNS cancer (glio) SF-295	40.6
Lung ca. A549	14.6	Brain (Amygdala) Pool	5.1
Lung ca. NCI-H526	4.8	Brain (cerebellum)	14.2
Lung ca. NCI-H23	18.0	Brain (fetal)	6.7
Lung ca. NCI-H460	8.4	Brain (Hippocampus) Pool	4.9
Lung ca. HOP-62	14.3	Cerebral Cortex Pool	5.1
Lung ca. NCI-H522	12.1	Brain (Substantia nigra) Pool	6.0
Liver	5.0	Brain (Thalamus) Pool	7.0
Fetal Liver	10.2	Brain (whole)	7.5
Liver ca. HepG2	70.7	Spinal Cord Pool	7.6
Kidney Pool	11.5	Adrenal Gland	9.3
Fetal Kidney	4.9	Pituitary gland Pool	3.3
Renal ca. 786-0	20.3	Salivary Gland	10.1
Renal ca. A498	6.1	Thyroid (female)	6.2
Renal ca. ACHN	6.7	Pancreatic ca. CAPAN2	21.9

Renal ca. UO-31	13.9	Pancreas Pool	10.7
-----------------	------	---------------	------

Table GC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4179, Run 173507602	Tissue Name	Rel. Exp.(%) Ag4179, Run 173507602
Secondary Th1 act	56.3	HUVEC IL-1beta	57.8
Secondary Th2 act	50.0	HUVEC IFN gamma	61.6
Secondary Tr1 act	3.8	HUVEC TNF alpha + IFN gamma	43.8
Secondary Th1 rest	24.3	HUVEC TNF alpha + IL4	50.0
Secondary Th2 rest	24.1	HUVEC IL-11	41.5
Secondary Tr1 rest	25.5	Lung Microvascular EC none	76.8
Primary Th1 act	28.3	Lung Microvascular EC TNFalpha + IL-1beta	58.2
Primary Th2 act	39.8	Microvascular Dermal EC none	54.3
Primary Tr1 act	33.7	Microvascular Dermal EC TNFalpha + IL-1beta	39.8
Primary Th1 rest	19.5	Bronchial epithelium TNFalpha + IL1beta	40.6
Primary Th2 rest	18.8	Small airway epithelium none	24.1
Primary Tr1 rest	24.0	Small airway epithelium TNFalpha + IL-1beta	43.5
CD45RA CD4 lymphocyte act	55.5	Coronary artery SMC rest	49.7
CD45RO CD4 lymphocyte act	35.4	Coronary artery SMC TNFalpha + IL-1beta	51.1
CD8 lymphocyte act	40.9	Astrocytes rest	46.7
Secondary CD8 lymphocyte rest	35.8	Astrocytes TNFalpha + IL-1beta	24.0
Secondary CD8 lymphocyte act	29.1	KU-812 (Basophil) rest	39.2
CD4 lymphocyte none	12.9	KU-812 (Basophil) PMA/ionomycin	54.7
2ry Th1/Th2/Tr1 _anti- CD95 CH11	34.9	CCD1106 (Keratinocytes) none	50.0
LAK cells rest	46.3	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	41.5
LAK cells IL-2	40.9	Liver cirrhosis	16.8

LAK cells IL-2+IL-12	43.5	NCI-H292 none	17.4
LAK cells IL-2+IFN gamma	43.8	NCI-H292 IL-4	23.2
LAK cells IL-2+IL-18	40.3	NCI-H292 IL-9	28.1
LAK cells PMA/ionomycin	10.6	NCI-H292 IL-13	30.6
NK Cells IL-2 rest	45.4	NCI-H292 IFN gamma	25.0
Two Way MLR 3 day	44.1	HPAEC none	44.1
Two Way MLR 5 day	38.2	HPAEC TNF alpha + IL-1 beta	49.7
Two Way MLR 7 day	40.6	Lung fibroblast none	39.8
PBMC rest	22.2	Lung fibroblast TNF alpha + IL-1 beta	51.1
PBMC PWM	34.6	Lung fibroblast IL-4	47.6
PBMC PHA-L	42.0	Lung fibroblast IL-9	50.7
Ramos (B cell) none	63.7	Lung fibroblast IL-13	40.6
Ramos (B cell) ionomycin	100.0	Lung fibroblast IFN gamma	55.1
B lymphocytes PWM	27.4	Dermal fibroblast CCD1070 rest	56.3
B lymphocytes CD40L and IL-4	48.0	Dermal fibroblast CCD1070 TNF alpha	59.5
EOL-1 dbcAMP	42.0	Dermal fibroblast CCD1070 IL-1 beta	45.1
EOL-1 dbcAMP PMA/ionomycin	31.0	Dermal fibroblast IFN gamma	31.6
Dendritic cells none	54.7	Dermal fibroblast IL-4	40.1
Dendritic cells LPS	39.2	Dermal Fibroblasts rest	30.6
Dendritic cells anti-CD40	61.1	Neutrophils TNFa+LPS	12.2
Monocytes rest	88.9	Neutrophils rest	36.1
Monocytes LPS	40.3	Colon	42.6
Macrophages rest	50.3	Lung	32.8
Macrophages LPS	37.1	Thymus	27.4
HUVEC none	41.5	Kidney	58.2

HUVEC starved	53.2		
---------------	------	--	--

Table GD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4179, Run 268695203	Tissue Name	Rel. Exp.(%) Ag4179, Run 268695203
Colon cancer 1	34.4	Bladder cancer NAT 2	0.0
Colon NAT 1	18.9	Bladder cancer NAT 3	0.2
Colon cancer 2	36.1	Bladder cancer NAT 4	7.9
Colon cancer NAT 2	17.9	Adenocarcinoma of the prostate 1	16.4
Colon cancer 3	76.3	Adenocarcinoma of the prostate 2	1.3
Colon cancer NAT 3	34.2	Adenocarcinoma of the prostate 3	6.3
Colon malignant cancer 4	100.0	Adenocarcinoma of the prostate 4	17.9
Colon normal adjacent tissue 4	12.9	Prostate cancer NAT 5	3.0
Lung cancer 1	15.4	Adenocarcinoma of the prostate 6	3.7
Lung NAT 1	1.1	Adenocarcinoma of the prostate 7	3.2
Lung cancer 2	25.0	Adenocarcinoma of the prostate 8	1.2
Lung NAT 2	1.5	Adenocarcinoma of the prostate 9	21.8
Squamous cell carcinoma 3	24.1	Prostate cancer NAT 10	0.7
Lung NAT 3	0.4	Kidney cancer 1	18.9
metastatic melanoma 1	15.6	Kidney NAT 1	6.8
Melanoma 2	3.8	Kidney cancer 2	51.1
Melanoma 3	1.6	Kidney NAT 2	19.1
metastatic melanoma 4	39.2	Kidney cancer 3	16.6
metastatic melanoma 5	36.6	Kidney NAT 3	3.8
Bladder cancer 1	0.5	Kidney cancer 4	21.9
Bladder cancer NAT 1	0.0	Kidney NAT 4	21.5
Bladder cancer 2	2.5		

General_screening_panel_v1.4 Summary: Ag4179 Highest expression of the CG100456-01 gene is detected in a breast cancer BT 549 cell line (CT=24). High expression of this gene is seen in cluster of breast, ovarian, colon, gastric, renal, lung, pancreatic, CNS, hepatic, prostate cancer cell lines and melanoma cell lines. Thus, therapeutic modulation of this gene product could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4179 Highest expression of the CG100456-01 gene is detected in ionomycin treated Ramos B cells (CT=28). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4179 Highest expression of the CG100456-01 gene is detected in malignant colon cancer (CT=26), with significant expression also seen in metastatic melanoma, colon, lung, bladder, prostate and kidney cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene could be as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

H. CG100466-01: Adenine Nucleotide Translocator 2 (ANT 2) (ADP/ATP Translocase 2)

Expression of gene CG100466-01 was assessed using the primer-probe set Ag4177, described in Table HA. Results of the RTQ-PCR runs are shown in Tables HB, HC and HD.

Table HA. Probe Name Ag4177

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tgtgggtaaagctgaagctg-3'	20	452	224
Probe	TET-5'-aggcctctgtgactgcctgggtaaga-3'-TAMRA	26	485	225
Reverse	5'-ttggtacaggcccttaatcc-3'	20	526	226

Table HB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4177, Run 215539616	Tissue Name	Rel. Exp.(%) Ag4177, Run 215539616
AD 1 Hippo	11.4	Control (Path) 3 Temporal Ctx	9.3
AD 2 Hippo	31.2	Control (Path) 4 Temporal Ctx	30.1
AD 3 Hippo	4.1	AD 1 Occipital Ctx	18.2
AD 4 Hippo	5.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	96.6	AD 3 Occipital Ctx	4.6
AD 6 Hippo	60.3	AD 4 Occipital Ctx	20.2
Control 2 Hippo	42.6	AD 5 Occipital Ctx	59.0
Control 4 Hippo	11.0	AD 6 Occipital Ctx	47.6
Control (Path) 3 Hippo	5.4	Control 1 Occipital Ctx	3.7
AD 1 Temporal Ctx	11.2	Control 2 Occipital Ctx	100.0
AD 2 Temporal Ctx	28.1	Control 3 Occipital Ctx	16.7

AD 3 Temporal Ctx	3.2	Control 4 Occipital Ctx	7.2
AD 4 Temporal Ctx	27.5	Control (Path) 1 Occipital Ctx	95.3
AD 5 Inf Temporal Ctx	93.3	Control (Path) 2 Occipital Ctx	8.6
AD 5 Sup Temporal Ctx	34.4	Control (Path) 3 Occipital Ctx	6.5
AD 6 Inf Temporal Ctx	42.9	Control (Path) 4 Occipital Ctx	15.3
AD 6 Sup Temporal Ctx	49.0	Control 1 Parietal Ctx	5.8
Control 1 Temporal Ctx	5.5	Control 2 Parietal Ctx	31.6
Control 2 Temporal Ctx	47.6	Control 3 Parietal Ctx	18.9
Control 3 Temporal Ctx	16.6	Control (Path) 1 Parietal Ctx	81.2
Control 3 Temporal Ctx	15.1	Control (Path) 2 Parietal Ctx	32.8
Control (Path) 1 Temporal Ctx	44.8	Control (Path) 3 Parietal Ctx	5.0
Control (Path) 2 Temporal Ctx	42.9	Control (Path) 4 Parietal Ctx	48.3

Table HC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4177, Run 221036658	Tissue Name	Rel. Exp.(%) Ag4177, Run 221036658
Adipose	2.4	Renal ca. TK-10	18.7
Melanoma* Hs688(A).T	5.2	Bladder	12.9
Melanoma* Hs688(B).T	6.5	Gastric ca. (liver met.) NCI-N87	17.8
Melanoma* M14	53.2	Gastric ca. KATO III	100.0
Melanoma* LOXIMVI	33.7	Colon ca. SW-948	19.5
Melanoma* SK-MEL-5	58.6	Colon ca. SW480	81.2
Squamous cell carcinoma SCC-4	21.5	Colon ca.* (SW480 met) SW620	57.0
Testis Pool	2.3	Colon ca. HT29	35.4
Prostate ca.* (bone met) PC-3	37.1	Colon ca. HCT-116	52.9

Prostate Pool	1.5	Colon ca. CaCo-2	23.7
Placenta	3.4	Colon cancer tissue	27.0
Uterus Pool	0.6	Colon ca. SW1116	8.5
Ovarian ca. OVCAR-3	34.2	Colon ca. Colo-205	17.7
Ovarian ca. SK-OV-3	27.0	Colon ca. SW-48	19.8
Ovarian ca. OVCAR-4	32.1	Colon Pool	3.3
Ovarian ca. OVCAR-5	19.3	Small Intestine Pool	1.9
Ovarian ca. IGROV-1	19.2	Stomach Pool	2.8
Ovarian ca. OVCAR-8	13.7	Bone Marrow Pool	1.0
Ovary	3.4	Fetal Heart	5.2
Breast ca. MCF-7	39.5	Heart Pool	1.7
Breast ca. MDA-MB-231	37.6	Lymph Node Pool	3.0
Breast ca. BT 549	60.7	Fetal Skeletal Muscle	1.3
Breast ca. T47D	34.4	Skeletal Muscle Pool	1.2
Breast ca. MDA-N	65.1	Spleen Pool	2.5
Breast Pool	3.2	Thymus Pool	3.5
Trachea	4.7	CNS cancer (glio/astro) U87-MG	25.2
Lung	0.9	CNS cancer (glio/astro) U-118-MG	26.2
Fetal Lung	4.8	CNS cancer (neuro;met) SK-N-AS	29.1
Lung ca. NCI-N417	13.8	CNS cancer (astro) SF-539	28.3
Lung ca. LX-1	46.3	CNS cancer (astro) SNB-75	43.8
Lung ca. NCI-H146	13.1	CNS cancer (glio) SNB-19	15.7
Lung ca. SHP-77	28.5	CNS cancer (glio) SF-295	18.8
Lung ca. A549	48.3	Brain (Amygdala) Pool	3.7
Lung ca. NCI-H526	10.2	Brain (cerebellum)	7.3
Lung ca. NCI-H23	24.7	Brain (fetal)	3.2
Lung ca. NCI-	13.7	Brain (Hippocampus) Pool	3.5

H460			
Lung ca. HOP-62	8.9	Cerebral Cortex Pool	5.1
Lung ca. NCI-H522	25.7	Brain (Substantia nigra) Pool	4.9
Liver	2.0	Brain (Thalamus) Pool	6.4
Fetal Liver	10.2	Brain (whole)	4.1
Liver ca. HepG2	23.0	Spinal Cord Pool	2.9
Kidney Pool	3.2	Adrenal Gland	3.9
Fetal Kidney	4.5	Pituitary gland Pool	0.8
Renal ca. 786-0	17.8	Salivary Gland	4.0
Renal ca. A498	6.5	Thyroid (female)	2.1
Renal ca. ACHN	14.5	Pancreatic ca. CAPAN2	29.3
Renal ca. UO-31	11.9	Pancreas Pool	4.8

Table HD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4177, Run 173333376	Tissue Name	Rel. Exp.(%) Ag4177, Run 173333376
Secondary Th1 act	72.2	HUVEC IL-1beta	54.3
Secondary Th2 act	51.8	HUVEC IFN gamma	34.6
Secondary Tr1 act	47.6	HUVEC TNF alpha + IFN gamma	23.5
Secondary Th1 rest	17.2	HUVEC TNF alpha + IL4	30.6
Secondary Th2 rest	17.4	HUVEC IL-11	22.4
Secondary Tr1 rest	17.8	Lung Microvascular EC none	39.0
Primary Th1 act	52.1	Lung Microvascular EC TNFalpha + IL-1beta	25.0
Primary Th2 act	60.3	Microvascular Dermal EC none	25.2
Primary Tr1 act	80.1	Microvascular Dermal EC TNFalpha + IL-1beta	24.0
Primary Th1 rest	25.5	Bronchial epithelium TNFalpha + IL1beta	40.6
Primary Th2 rest	19.6	Small airway epithelium none	16.3
Primary Tr1 rest	37.6	Small airway epithelium TNFalpha + IL-1beta	44.8
CD45RA CD4 lymphocyte act	61.1	Coronary artery SMC rest	16.5
CD45RO CD4 lymphocyte act	74.7	Coronary artery SMC TNFalpha + IL-1beta	16.6

CD8 lymphocyte act	96.6	Astrocytes rest	11.2
Secondary CD8 lymphocyte rest	57.0	Astrocytes TNFalpha + IL-1beta	11.6
Secondary CD8 lymphocyte act	38.4	KU-812 (Basophil) rest	27.2
CD4 lymphocyte none	6.2	KU-812 (Basophil) PMA/ionomycin	41.2
2ry Th1/Th2/Tr1_anti-CD95 CH11	25.0	CCD1106 (Keratinocytes) none	52.1
LAK cells rest	31.6	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	57.0
LAK cells IL-2	38.4	Liver cirrhosis	7.0
LAK cells IL-2+IL-12	25.9	NCI-H292 none	56.6
LAK cells IL-2+IFN gamma	32.8	NCI-H292 IL-4	61.1
LAK cells IL-2+IL-18	36.3	NCI-H292 IL-9	100.0
LAK cells PMA/ionomycin	56.3	NCI-H292 IL-13	51.1
NK Cells IL-2 rest	45.4	NCI-H292 IFN gamma	4.7
Two Way MLR 3 day	25.0	HPAEC none	24.8
Two Way MLR 5 day	49.3	HPAEC TNF alpha + IL-1 beta	29.7
Two Way MLR 7 day	42.6	Lung fibroblast none	14.8
PBMC rest	11.0	Lung fibroblast TNF alpha + IL-1 beta	11.4
PBMC PWM	50.0	Lung fibroblast IL-4	17.2
PBMC PHA-L	54.3	Lung fibroblast IL-9	19.5
Ramos (B cell) none	67.4	Lung fibroblast IL-13	15.3
Ramos (B cell) ionomycin	82.4	Lung fibroblast IFN gamma	21.2
B lymphocytes PWM	48.6	Dermal fibroblast CCD1070 rest	33.7
B lymphocytes CD40L and IL-4	54.0	Dermal fibroblast CCD1070 TNF alpha	57.4
EOL-1 dbcAMP	44.1	Dermal fibroblast CCD1070 IL-1 beta	25.2

EOL-1 dbcAMP PMA/ionomycin	31.9	Dermal fibroblast IFN gamma	20.3
Dendritic cells none	45.1	Dermal fibroblast IL-4	28.3
Dendritic cells LPS	31.2	Dermal Fibroblasts rest	15.1
Dendritic cells anti-CD40	36.1	Neutrophils TNFa+LPS	3.0
Monocytes rest	24.1	Neutrophils rest	4.3
Monocytes LPS	13.4	Colon	29.9
Macrophages rest	50.3	Lung	12.9
Macrophages LPS	15.4	Thymus	29.1
HUVEC none	32.8	Kidney	51.4
HUVEC starved	37.9		

CNS_neurodegeneration_v1.0 Summary: Ag4177 This panel confirms the expression of the CG100466-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4177 Highest expression of the CG100466-01 gene is detected in a gastric cancer KATO III cell line (CT=22). High expression of this gene is seen in cluster of breast, ovarian, colon, gastric, renal, lung, pancreatic, CNS, hepatic, prostate cancer cell lines and melanoma cell lines. Thus, therapeutic modulation of this gene product could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

The CG100466-01 gene codes for adenine nucleotide translocator (ANT 2) homologue. Dysfunctioning of the ANT2 have been shown to induce myopathies in mouse and in humans (Fiore et al., 2001, Clin Chim Acta 311(2):125-35, PMID: 11566172). ANT has a role in mtDNA maintenance and mutation in ANT has been

implicated in autosomal dominant progressive external ophthalmoplegia and other mitochondrial diseases (Kaukonen et al., 2000, Science 289(5480):782-5, PMID: 10926541). Mice deficient in the heart/muscle specific isoform of the ANT1 exhibit many of the hallmarks of human oxidative phosphorylation (OXPHOS) disease, including a dramatic proliferation of skeletal muscle mitochondria (Murdoch et al., 1999, J Biol Chem 274(20):14429-33, PMID: 10318868). Therefore, therapeutic modulation of the ANT protein encoded by the CG100466-01 gene through the use of small molecule drug could be useful in the treatment mitochondria related diseases including autosomal dominant progressive external ophthalmoplegia and cancers.

10 In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

15 **Panel 4.1D Summary:** Ag4167 Highest expression of the CG100466-01 gene is detected in IL-9 treated NCI-H292 cell line (CT=25). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

30 **General oncology screening panel_v_2.4 Summary:** Ag4177 Results from one experiment with the CG100466-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

I. CG100609-01: Glutathione S-Transferase

Expression of gene CG100609-01 was assessed using the primer-probe set Ag4182, described in Table IA. Results of the RTQ-PCR runs are shown in Tables IB, IC and ID.

Table IA. Probe Name Ag4182

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctctacatggacctgctgtca-3'	21	73	227
Probe	TET-5'-ccgtgccgtctacatcttctogaag-3'-TAMRA	25	102	228
Reverse	5'-agttgaactggatgtcatgctt-3'	22	127	229

Table IB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4182, Run 215539692	Tissue Name	Rel. Exp.(%) Ag4182, Run 215539692
AD 1 Hippo	0.0	Control (Path) 3 Temporal Ctx	0.0
AD 2 Hippo	0.0	Control (Path) 4 Temporal Ctx	23.8
AD 3 Hippo	0.0	AD 1 Occipital Ctx	0.0
AD 4 Hippo	0.0	AD 2 Occipital Ctx (Missing)	2.0
AD 5 hippo	0.0	AD 3 Occipital Ctx	0.0
AD 6 Hippo	32.3	AD 4 Occipital Ctx	0.0
Control 2 Hippo	0.0	AD 5 Occipital Ctx	0.0
Control 4 Hippo	22.7	AD 6 Occipital Ctx	0.0
Control (Path) 3 Hippo	0.0	Control 1 Occipital Ctx	0.0
AD 1 Temporal Ctx	78.5	Control 2 Occipital Ctx	0.0
AD 2 Temporal Ctx	0.0	Control 3 Occipital Ctx	0.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	22.2
AD 4 Temporal Ctx	63.3	Control (Path) 1 Occipital Ctx	25.0
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	0.0
AD 5 SupTemporal Ctx	0.0	Control (Path) 3 Occipital Ctx	0.0

AD 6 Inf Temporal Ctx	23.7	Control (Path) 4 Occipital Ctx	0.0
AD 6 Sup Temporal Ctx	27.4	Control 1 Parietal Ctx	21.2
Control 1 Temporal Ctx	0.0	Control 2 Parietal Ctx	0.0
Control 2 Temporal Ctx	0.0	Control 3 Parietal Ctx	0.0
Control 3 Temporal Ctx	0.0	Control (Path) 1 Parietal Ctx	0.0
Control 4 Temporal Ctx	0.0	Control (Path) 2 Parietal Ctx	0.0
Control (Path) 1 Temporal Ctx	0.0	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	0.0	Control (Path) 4 Parietal Ctx	0.0

Table IC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4182, Run 221118504	Tissue Name	Rel. Exp.(%) Ag4182, Run 221118504
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.6
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.8
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	1.5
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	100.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	2.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	9.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.6	Colon ca. SW-48	0.0

Ovarian ca. OVCAR-4	0.6	Colon Pool	0.4
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.7
Ovarian ca. IGROV-1	1.0	Stomach Pool	0.5
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.4
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.6	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	1.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.4
Breast Pool	0.5	Thymus Pool	0.0
Trachea	1.6	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.0	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	5.4	CNS cancer (astro) SNB-75	0.8
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.6
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	2.3
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	0.3
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	0.7
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.8
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.0	Brain (whole)	0.0
Liver ca. HepG2	0.7	Spinal Cord Pool	0.0

Kidney Pool	0.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.7	Salivary Gland	0.8
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.6
Renal ca. UO-31	0.0	Pancreas Pool	0.5

Table ID. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4182, Run 173607859	Tissue Name	Rel. Exp.(%) Ag4182, Run 173607859
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.9	Lung Microvascular EC none	0.0
Primary Th1 act	0.9	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.6
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry	0.0	CCD1106 (Keratinocytes) none	0.0

Th1/Th2/Tr1_anti- CD95 CH11			
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	1.7
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.5	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	1.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.8	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	1.5
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	1.5
B lymphocytes CD40L and IL-4	1.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.7
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	1.1
Monocytes rest	0.0	Neutrophils rest	1.6

Monocytes LPS	0.0	Colon	0.8
Macrophages rest	0.0	Lung	0.8
Macrophages LPS	0.0	Thymus	7.9
HUVEC none	0.7	Kidney	100.0
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4182 This panel confirms the expression of the CG100609-01 gene at very low levels in the brain in an independent group of individuals. This gene is found to be upregulated in the temporal cortex of Alzheimer's disease patients. Therefore, therapeutic modulation of the expression or function of this gene may decrease neuronal death and be of use in the treatment of this disease.

General_screening_panel_v1.4 Summary: Ag4182 Highest expression of the CG100609-01 gene is detected in testis (CT=29). Thus, expression of this gene can be used to distinguish testis from other samples in this panel. In addition, therapeutic modulation of this gene can be useful in the treatment of testis related diseases such as fertility and hypogonadism.

Low levels of expression of this gene is also associated with a CNS cancer, colon cancer, lung cancer, and an ovarian cancer cell lines. Therefore, therapeutic modulation of this gene can be beneficial in the treatments of these cancers.

Panel 4.1D Summary: Ag4182 Highest expression of the CG100609-01 gene is detected exclusively in kidney (CT=30.6). In addition low expression of this gene is also seen in thymus. Thus, expression of this gene can be used to distinguish these two samples from other samples in this panel. Furthermore, small molecule therapies designed with the protein encoded for by this gene could modulate kidney function and be important in the treatment of inflammatory or autoimmune diseases that affect the kidney, including lupus and glomerulonephritis

Panel CNS_1 Summary: Ag4182 Expression of the CG100609-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4182 Expression of the CG100609-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

J. CG100710-01: AAA (ATPase Associated with Various Activities)

Expression of gene CG100710-01 was assessed using the primer-probe set Ag4249, described in Table JA.

Table JA. Probe Name Ag4249

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - cacaggagcggctgtca - 3'	17	213	230
Probe	TET-5' - tgccagccctgagcaagtgcc - 3' - TAMRA	21	235	231
Reverse	5' - ctgcagagcacagcactca - 3'	19	262	232

5 **CNS_neurodegeneration_v1.0 Summary:** Ag4249 Expression of the CG100710-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4249 Expression of the CG100710-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

10 **Panel 4.1D Summary:** Ag4249 Expression of the CG100710-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4249 Expression of the CG100710-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

15

K. CG100730-01: Exoribonuclease

Expression of gene CG100730-01 was assessed using the primer-probe set Ag4187, described in Table KA. Results of the RTQ-PCR runs are shown in Tables KB, KC, KD and KE.

20

Table KA. Probe Name Ag4187

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gcaaatggctctgctgtaatac - 3'	22	201	233
Probe	TET-5' - taatggccacagccgtcagtaaaaca - 3' - TAMRA	26	241	234
Reverse	5' - taaactgggaaggggaaggt - 3'	20	269	235

Table KB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4187, Run 215539701	Tissue Name	Rel. Exp.(%) Ag4187, Run 215539701
AD 1 Hippo	0.0	Control (Path) 3 Temporal Ctx	7.5

AD 2 Hippo	0.0	Control (Path) 4 Temporal Ctx	0.4
AD 3 Hippo	0.0	AD 1 Occipital Ctx	0.0
AD 4 Hippo	0.0	AD 2 Occipital Ctx (Missing)	100.0
AD 5 Hippo	1.9	AD 3 Occipital Ctx	0.0
AD 6 Hippo	3.0	AD 4 Occipital Ctx	1.3
Control 2 Hippo	0.0	AD 5 Occipital Ctx	1.1
Control 4 Hippo	0.0	AD 6 Occipital Ctx	2.6
Control (Path) 3 Hippo	0.3	Control 1 Occipital Ctx	0.0
AD 1 Temporal Ctx	0.0	Control 2 Occipital Ctx	0.0
AD 2 Temporal Ctx	0.0	Control 3 Occipital Ctx	0.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	0.0
AD 4 Temporal Ctx	0.9	Control (Path) 1 Occipital Ctx	1.2
AD 5 Inf Temporal Ctx	2.9	Control (Path) 2 Occipital Ctx	0.0
AD 5 Sup Temporal Ctx	2.1	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	9.9	Control (Path) 4 Occipital Ctx	1.0
AD 6 Sup Temporal Ctx	7.5	Control 1 Parietal Ctx	0.0
Control 1 Temporal Ctx	0.0	Control 2 Parietal Ctx	12.8
Control 2 Temporal Ctx	0.0	Control 3 Parietal Ctx	0.0
Control 3 Temporal Ctx	0.0	Control (Path) 1 Parietal Ctx	0.1
Control 3 Temporal Ctx	0.0	Control (Path) 2 Parietal Ctx	5.8
Control (Path) 1 Temporal Ctx	0.0	Control (Path) 3 Parietal Ctx	6.7
Control (Path) 2 Temporal Ctx	0.0	Control (Path) 4 Parietal Ctx	0.4

Table KC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4187, Run 221154082	Tissue Name	Rel. Exp.(%) Ag4187, Run 221154082
-------------	------------------------------------	-------------	------------------------------------

Adipose	13.2	Renal ca. TK-10	41.5
Melanoma* Hs688(A).T	2.2	Bladder	16.4
Melanoma* Hs688(B).T	5.7	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	4.4	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	9.7
Testis Pool	10.7	Colon ca. HT29	1.0
Prostate ca.* (bone met) PC-3	9.9	Colon ca. HCT-116	5.2
Prostate Pool	19.8	Colon ca. CaCo-2	0.0
Placenta	4.1	Colon cancer tissue	18.6
Uterus Pool	6.1	Colon ca. SW1116	1.0
Ovarian ca. OVCAR-3	6.9	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	31.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	9.6
Ovarian ca. OVCAR-5	55.1	Small Intestine Pool	33.0
Ovarian ca. IGROV-1	1.6	Stomach Pool	21.9
Ovarian ca. OVCAR-8	12.9	Bone Marrow Pool	0.8
Ovary	10.5	Fetal Heart	0.0
Breast ca. MCF-7	17.3	Heart Pool	14.7
Breast ca. MDA-MB-231	9.1	Lymph Node Pool	16.3
Breast ca. BT 549	24.8	Fetal Skeletal Muscle	7.2
Breast ca. T47D	75.8	Skeletal Muscle Pool	24.3
Breast ca. MDA-N	0.8	Spleen Pool	26.4
Breast Pool	23.3	Thymus Pool	26.2
Trachea	25.2	CNS cancer (glio/astro) U87-MG	8.5
Lung	3.4	CNS cancer (glio/astro) U-	3.7

		118-MG	
Fetal Lung	0.0	CNS cancer (neuro;met) SK-N-AS	0.7
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	4.9
Lung ca. LX-1	29.3	CNS cancer (astro) SNB-75	100.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	5.8
Lung ca. SHP-77	2.9	CNS cancer (glio) SF-295	76.3
Lung ca. A549	0.0	Brain (Amygdala) Pool	13.8
Lung ca. NCI-H526	1.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	49.7
Lung ca. NCI-H460	41.8	Brain (Hippocampus) Pool	36.9
Lung ca. HOP-62	25.7	Cerebral Cortex Pool	18.9
Lung ca. NCI-H522	2.4	Brain (Substantia nigra) Pool	7.7
Liver	0.0	Brain (Thalamus) Pool	34.2
Fetal Liver	3.7	Brain (whole)	15.8
Liver ca. HepG2	4.7	Spinal Cord Pool	15.5
Kidney Pool	45.7	Adrenal Gland	2.0
Fetal Kidney	0.0	Pituitary gland Pool	1.9
Renal ca. 786-0	2.3	Salivary Gland	1.3
Renal ca. A498	3.6	Thyroid (female)	2.5
Renal ca. ACHN	21.0	Pancreatic ca. CAPAN2	92.7
Renal ca. UO-31	3.9	Pancreas Pool	10.8

Table KD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4187, Run 182086758	Tissue Name	Rel. Exp.(%) Ag4187, Run 182086758
Secondary Th1 act	2.0	HUVEC IL-1beta	0.0
Secondary Th2 act	1.8	HUVEC IFN gamma	2.4
Secondary Tr1 act	3.0	HUVEC TNF alpha + IFN gamma	0.5
Secondary Th1 rest	3.1	HUVEC TNF alpha + IL4	0.2
Secondary Th2 rest	0.0	HUVEC IL-11	4.3
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.4
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha	1.8

		+ IL-1beta	
Primary Th2 act	0.0	Microvascular Dermal EC none	1.8
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	3.6
Primary Th1 rest	7.2	Bronchial epithelium TNFalpha + IL1beta	1.7
Primary Th2 rest	4.0	Small airway epithelium none	0.0
Primary Tr1 rest	5.5	Small airway epithelium TNFalpha + IL-1beta	1.3
CD45RA CD4 lymphocyte act	4.8	Coronary artery SMC rest	0.4
CD45RO CD4 lymphocyte act	7.6	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	6.7	Astrocytes rest	0.8
Secondary CD8 lymphocyte rest	3.4	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	5.3	KU-812 (Basophil) rest	1.0
CD4 lymphocyte none	6.6	KU-812 (Basophil) PMA/ionomycin	16.4
2ry Th1/Th2/Tr1 _anti- CD95 CH11	4.5	CCD1106 (Keratinocytes) none	2.3
LAK cells rest	10.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	1.9
LAK cells IL-2	5.0	Liver cirrhosis	0.7
LAK cells IL- 2+IL-12	3.1	NCI-H292 none	0.0
LAK cells IL- 2+IFN gamma	6.7	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	7.7	NCI-H292 IL-9	0.2
LAK cells PMA/ionomycin	3.8	NCI-H292 IL-13	0.5
NK Cells IL-2 rest	3.4	NCI-H292 IFN gamma	0.4
Two Way MLR 3 day	15.9	HPAEC none	1.4
Two Way MLR 5 day	3.1	HPAEC TNF alpha + IL-1 beta	3.2
Two Way MLR 7 day	0.5	Lung fibroblast none	7.0
PBMC rest	1.1	Lung fibroblast TNF alpha + IL-1	1.8

		beta	
PBMC PWM	4.1	Lung fibroblast IL-4	1.7
PBMC PHA-L	3.3	Lung fibroblast IL-9	2.8
Ramos (B cell) none	2.8	Lung fibroblast IL-13	0.3
Ramos (B cell) ionomycin	0.3	Lung fibroblast IFN gamma	0.5
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	3.4
B lymphocytes CD40L and IL-4	0.3	Dermal fibroblast CCD1070 TNF alpha	6.0
EOL-1 dbcAMP	9.9	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	2.7	Dermal fibroblast IFN gamma	0.9
Dendritic cells none	10.7	Dermal fibroblast IL-4	1.9
Dendritic cells LPS	4.7	Dermal Fibroblasts rest	1.1
Dendritic cells anti-CD40	8.3	Neutrophils TNFa+LPS	4.3
Monocytes rest	19.2	Neutrophils rest	7.4
Monocytes LPS	2.7	Colon	0.0
Macrophages rest	8.5	Lung	3.7
Macrophages LPS	0.0	Thymus	17.9
HUVEC none	0.8	Kidney	100.0
HUVEC starved	1.0		

Table KE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4187, Run 268689531	Tissue Name	Rel. Exp.(%) Ag4187, Run 268689531
Colon cancer 1	7.8	Bladder cancer NAT 2	1.5
Colon cancer NAT 1	0.0	Bladder cancer NAT 3	2.4
Colon cancer 2	0.0	Bladder cancer NAT 4	1.6
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	86.5
Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	2.1
Colon cancer NAT 3	8.8	Adenocarcinoma of the prostate 3	0.0
Colon malignant	2.3	Adenocarcinoma of the	6.2

cancer 4		prostate 4	
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	15.2
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	14.7
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	1.9
Lung NAT 2	3.9	Adenocarcinoma of the prostate 9	18.0
Squamous cell carcinoma 3	0.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	3.3
metastatic melanoma 1	26.8	Kidney NAT 1	13.4
Melanoma 2	0.0	Kidney cancer 2	71.2
Melanoma 3	6.8	Kidney NAT 2	23.7
metastatic melanoma 4	21.9	Kidney cancer 3	79.0
metastatic melanoma 5	95.3	Kidney NAT 3	7.5
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	3.5		

CNS_neurodegeneration_v1.0 Summary: Ag4187 Results from one experiment with the CG100730-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4187 Highest expression of the CG100730-01 gene is detected in CNS cancer SNB-75 cell line (CT=32). In addition high expression of this gene is seen in pancreatic cancer, CNS cancer, renal, lung, breast and ovarian cancer cell line. Therefore, therapeutic modulation of this gene through the use of small molecule drugs can be useful in the treatment of these cancers.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, and cerebral cortex. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4187 Highest expression of the CG100730-01 gene is detected in kidney (CT=32.7). Thus, expression of this gene can be used to distinguish kidney samples from other samples in this panel. Furthermore, small molecule therapies designed with the protein encoded for by this gene could modulate kidney function and be important in the treatment of inflammatory or autoimmune diseases that affect the kidney, including lupus and glomerulonephritis.

In addition, low expression of this gene is seen in resting primary Th1 and Tr1 cells, activated CD45RO CD4 lymphocyte, 3 day Two Way MLR, eosinophils, monocytes, macrophages, dendritic cells, ionomycin treated basophils and thymus. Therefore, therapeutic modulation of this gene may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4187 Highest expression of the CG100730-01 gene is detected in lung cancer sample (CT=33.7), with significant expression also seen in metastatic melanoma, lung, prostate and kidney cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene could be as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

L. CG100819-01: Polynucleotide Phosphorylase

Expression of gene CG100819-01 was assessed using the primer-probe set Ag4195, described in Table LA. Results of the RTQ-PCR runs are shown in Tables LB, LC and LD.

Table LA. Probe Name Ag4195

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-atcggtcaattagaccgctctt-3'	22	430	236
Probe	TET-5'-tccagctggctacttctatgatacacagg-3'-TAMRA	29	452	237
Reverse	5'-acatcaggctcatttacaccat-3'	22	505	238

Table LB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4195, Run 221157738	Tissue Name	Rel. Exp.(%) Ag4195, Run 221157738
Adipose	3.4	Renal ca. TK-10	16.8
Melanoma* Hs688(A).T	5.7	Bladder	13.3
Melanoma* Hs688(B).T	5.6	Gastric ca. (liver met.) NCI- N87	100.0
Melanoma* M14	16.8	Gastric ca. KATO III	58.2
Melanoma* LOXIMVI	25.3	Colon ca. SW-948	7.1
Melanoma* SK- MEL-5	25.9	Colon ca. SW480	35.8
Squamous cell carcinoma SCC-4	10.1	Colon ca.* (SW480 met) SW620	19.2
Testis Pool	5.1	Colon ca. HT29	9.0
Prostate ca.* (bone met) PC-3	20.4	Colon ca. HCT-116	27.5
Prostate Pool	2.1	Colon ca. CaCo-2	30.8
Placenta	0.6	Colon cancer tissue	8.8
Uterus Pool	2.3	Colon ca. SW1116	3.0
Ovarian ca. OVCAR-3	10.6	Colon ca. Colo-205	4.7
Ovarian ca. SK- OV-3	17.6	Colon ca. SW-48	7.3
Ovarian ca. OVCAR-4	5.8	Colon Pool	4.9
Ovarian ca. OVCAR-5	12.8	Small Intestine Pool	3.2
Ovarian ca. IGROV-1	5.8	Stomach Pool	2.7
Ovarian ca. OVCAR-8	3.1	Bone Marrow Pool	1.8
Ovary	2.5	Fetal Heart	3.7
Breast ca. MCF-7	14.3	Heart Pool	3.4
Breast ca. MDA- MB-231	15.1	Lymph Node Pool	6.3
Breast ca. BT 549	26.8	Fetal Skeletal Muscle	1.8
Breast ca. T47D	21.0	Skeletal Muscle Pool	7.7
Breast ca. MDA-N	11.3	Spleen Pool	5.0
Breast Pool	5.6	Thymus Pool	4.4
Trachea	3.1	CNS cancer (glio/astro) U87-	14.0

		MG	
Lung	0.9	CNS cancer (glio/astro) U-118-MG	21.0
Fetal Lung	8.6	CNS cancer (neuro;met) SK-N-AS	22.1
Lung ca. NCI-N417	3.5	CNS cancer (astro) SF-539	3.1
Lung ca. LX-1	20.0	CNS cancer (astro) SNB-75	12.4
Lung ca. NCI-H146	6.7	CNS cancer (glio) SNB-19	5.3
Lung ca. SHP-77	39.0	CNS cancer (glio) SF-295	20.4
Lung ca. A549	23.0	Brain (Amygdala) Pool	3.9
Lung ca. NCI-H526	4.4	Brain (cerebellum)	2.2
Lung ca. NCI-H23	21.3	Brain (fetal)	3.8
Lung ca. NCI-H460	16.8	Brain (Hippocampus) Pool	3.7
Lung ca. HOP-62	6.5	Cerebral Cortex Pool	5.0
Lung ca. NCI-H522	12.3	Brain (Substantia nigra) Pool	3.4
Liver	0.4	Brain (Thalamus) Pool	7.2
Fetal Liver	6.3	Brain (whole)	3.5
Liver ca. HepG2	8.2	Spinal Cord Pool	4.2
Kidney Pool	6.2	Adrenal Gland	2.8
Fetal Kidney	5.4	Pituitary gland Pool	1.3
Renal ca. 786-0	5.3	Salivary Gland	0.6
Renal ca. A498	3.1	Thyroid (female)	1.8
Renal ca. ACHN	6.4	Pancreatic ca. CAPAN2	8.5
Renal ca. UO-31	9.9	Pancreas Pool	5.4

Table LC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4195, Run 174255656	Tissue Name	Rel. Exp.(%) Ag4195, Run 174255656
Secondary Th1 act	54.3	HUVEC IL-1beta	20.6
Secondary Th2 act	100.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	28.7	HUVEC TNF alpha + IFN gamma	26.4
Secondary Th1 rest	7.5	HUVEC TNF alpha + IL4	15.6
Secondary Th2 rest	6.6	HUVEC IL-11	11.4
Secondary Tr1 rest	7.5	Lung Microvascular EC none	20.7
Primary Th1 act	57.0	Lung Microvascular EC TNFalpha + IL-1beta	15.9
Primary Th2 act	47.0	Microvascular Dermal EC none	19.8
Primary Tr1 act	51.4	Microvascular Dermal EC	12.0

		TNFalpha + IL-1beta	
Primary Th1 rest	5.2	Bronchial epithelium TNFalpha + IL1beta	10.5
Primary Th2 rest	3.0	Small airway epithelium none	5.7
Primary Tr1 rest	5.4	Small airway epithelium TNFalpha + IL-1beta	12.4
CD45RA CD4 lymphocyte act	46.7	Coronary artery SMC rest	11.7
CD45RO CD4 lymphocyte act	62.0	Coronary artery SMC TNFalpha + IL-1beta	10.0
CD8 lymphocyte act	37.1	Astrocytes rest	4.8
Secondary CD8 lymphocyte rest	41.2	Astrocytes TNFalpha + IL-1beta	6.1
Secondary CD8 lymphocyte act	14.8	KU-812 (Basophil) rest	29.1
CD4 lymphocyte none	4.0	KU-812 (Basophil) PMA/ionomycin	49.3
2ry Th1/Th2/Tr1_anti-CD95 CH11	8.1	CCD1106 (Keratinocytes) none	24.0
LAK cells rest	12.3	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	24.5
LAK cells IL-2	25.5	Liver cirrhosis	2.3
LAK cells IL-2+IL-12	28.9	NCI-H292 none	13.1
LAK cells IL-2+IFN gamma	25.3	NCI-H292 IL-4	29.5
LAK cells IL-2+ IL-18	33.4	NCI-H292 IL-9	30.1
LAK cells PMA/ionomycin	31.2	NCI-H292 IL-13	31.6
NK Cells IL-2 rest	24.7	NCI-H292 IFN gamma	47.6
Two Way MLR 3 day	15.5	HPAEC none	14.2
Two Way MLR 5 day	24.1	HPAEC TNF alpha + IL-1 beta	27.9
Two Way MLR 7 day	13.5	Lung fibroblast none	13.1
PBMC rest	4.6	Lung fibroblast TNF alpha + IL-1 beta	22.5
PBMC PWM	29.5	Lung fibroblast IL-4	16.2
PBMC PHA-L	18.4	Lung fibroblast IL-9	21.2
Ramos (B cell) none	50.7	Lung fibroblast IL-13	15.8

Ramos (B cell) ionomycin	56.3	Lung fibroblast IFN gamma	20.2
B lymphocytes PWM	37.4	Dermal fibroblast CCD1070 rest	23.7
B lymphocytes CD40L and IL-4	13.7	Dermal fibroblast CCD1070 TNF alpha	26.8
EOL-1 dbcAMP	22.7	Dermal fibroblast CCD1070 IL-1 beta	16.0
EOL-1 dbcAMP PMA/ionomycin	7.6	Dermal fibroblast IFN gamma	7.0
Dendritic cells none	7.6	Dermal fibroblast IL-4	10.6
Dendritic cells LPS	27.9	Dermal Fibroblasts rest	5.7
Dendritic cells anti-CD40	10.5	Neutrophils TNFa+LPS	0.9
Monocytes rest	8.2	Neutrophils rest	2.5
Monocytes LPS	41.5	Colon	3.1
Macrophages rest	9.9	Lung	6.8
Macrophages LPS	30.4	Thymus	10.4
HUVEC none	12.5	Kidney	7.4
HUVEC starved	14.0		

Table LD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4195, Run 268689535	Tissue Name	Rel. Exp.(%) Ag4195, Run 268689535
Colon cancer 1	23.2	Bladder cancer NAT 2	0.7
Colon cancer NAT 1	7.3	Bladder cancer NAT 3	1.6
Colon cancer 2	45.4	Bladder cancer NAT 4	4.0
Colon cancer NAT 2	19.6	Adenocarcinoma of the prostate 1	26.2
Colon cancer 3	95.9	Adenocarcinoma of the prostate 2	2.0
Colon cancer NAT 3	18.7	Adenocarcinoma of the prostate 3	15.5
Colon malignant cancer 4	72.7	Adenocarcinoma of the prostate 4	30.6
Colon normal adjacent tissue 4	8.0	Prostate cancer NAT 5	3.3
Lung cancer 1	18.0	Adenocarcinoma of the prostate 6	4.6
Lung NAT 1	2.2	Adenocarcinoma of the prostate 7	5.8
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	1.8

Lung NAT 2	3.4	Adenocarcinoma of the prostate 9	18.3
Squamous cell carcinoma 3	39.2	Prostate cancer NAT 10	2.1
Lung NAT 3	0.6	Kidney cancer 1	17.1
metastatic melanoma 1	22.7	Kidney NAT 1	7.7
Melanoma 2	2.0	Kidney cancer 2	38.4
Melanoma 3	2.5	Kidney NAT 2	15.5
metastatic melanoma 4	31.0	Kidney cancer 3	10.2
metastatic melanoma 5	44.4	Kidney NAT 3	7.2
Bladder cancer 1	2.5	Kidney cancer 4	18.7
Bladder cancer NAT 1	0.0	Kidney NAT 4	7.7
Bladder cancer 2	3.8		

CNS_neurodegeneration_v1.0 Summary: Ag4195 Results from one experiment with the CG100819-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4195 Highest expression of the CG100819-01 gene is detected in a gastric cancer NCI-N87 cell line (CT=25.3). High expression of this gene is seen in cluster of CNS cancer, colon, gastric, renal, lung, breast, ovarian, pancreatic, prostate, squamous cell carcinoma cell lines and melanoma. Therefore, therapeutic modulation of this gene could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, expression of this gene is higher in fetal lung and liver (CTs=29) as compared to the corresponding adult tissues (CTs=32-33). Therefore, expression of this gene could be useful in distinguishing the fetal lung and liver from the corresponding adult tissues.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in

central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4195 Highest expression of the CG100819-01 gene is detected in activated secondary Th2 cells (CT=27). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4195 Highest expression of the CG100819-01 gene is detected in lung cancer (CT=28.7), with significant expression also seen in metastatic melanoma, colon, lung, prostate and kidney cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene could be as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

25

M. CG100872-01: Protein-Arginine Deiminase

Expression of gene CG100872-01 was assessed using the primer-probe sets Ag4197 and Ag4299, described in Tables MA and MB. Results of the RTQ-PCR runs are shown in Table MC.

30

Table MA. Probe Name Ag4197

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtcccagatgactctgaatgtc-3'	22	539	239

Probe	TET-5'-caaggccccagctgtatcttaaagaa-3'-TAMRA	26	561	240
Reverse	5'-ccttgagggtatggaggactag-3'	22	594	241

Table MB. Probe Name Ag4299

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtcccagatgactctgaatgtc-3'	22	539	242
Probe	TET-5'-caaggccccagctgtatcttaaagaa-3'-TAMRA	26	561	243
Reverse	5'-ttcgactcttccttgagggtat-3'	22	604	244

Table MC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4197, Run 174255682	Rel. Exp.(%) Ag4299, Run 183712613	Tissue Name	Rel. Exp.(%) Ag4197, Run 174255682	Rel. Exp.(%) Ag4299, Run 183712613
Secondary Th1 act	0.0	0.0	HUVEC IL-1beta	0.0	0.0
Secondary Th2 act	0.0	0.0	HUVEC IFN gamma	0.0	0.0
Secondary Tr1 act	0.5	29.1	HUVEC TNF alpha + IFN gamma	0.0	0.0
Secondary Th1 rest	0.0	0.0	HUVEC TNF alpha + IL4	0.0	0.0
Secondary Th2 rest	0.0	0.0	HUVEC IL-11	0.0	0.0
Secondary Tr1 rest	0.0	0.0	Lung Microvascular EC none	0.0	0.0
Primary Th1 act	0.0	0.0	Lung Microvascular EC TNFalpha + IL- 1beta	0.0	0.0
Primary Th2 act	0.0	0.0	Microvascular Dermal EC none	0.0	0.0
Primary Tr1 act	0.0	0.0	Microvascular Dermal EC TNFalpha + IL- 1beta	0.0	0.0
Primary Th1 rest	0.0	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0	0.0
Primary Th2 rest	0.0	0.0	Small airway epithelium none	0.0	0.0
Primary Tr1 rest	0.0	0.0	Small airway epithelium TNFalpha + IL- 1beta	0.0	0.0
CD45RA CD4 lymphocyte act	0.0	0.0	Coronary artery SMC rest	0.0	0.0

CD45RO CD4 lymphocyte act	0.0	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0	0.0
CD8 lymphocyte act	0.0	0.0	Astrocytes rest	0.0	0.0
Secondary CD8 lymphocyte rest	0.0	0.0	Astrocytes TNFalpha + IL-1beta	0.0	0.0
Secondary CD8 lymphocyte act	0.0	0.0	KU-812 (Basophil) rest	0.0	0.0
CD4 lymphocyte none	0.0	0.0	KU-812 (Basophil) PMA/ionomycin	0.0	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	0.0	CCD1106 (Keratinocytes) none	0.0	0.0
LAK cells rest	0.0	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0	0.0
LAK cells IL-2	0.0	0.0	Liver cirrhosis	0.0	0.0
LAK cells IL-2+IL-12	0.0	0.0	NCI-H292 none	0.0	0.0
LAK cells IL-2+IFN gamma	0.0	0.0	NCI-H292 IL-4	0.0	0.0
LAK cells IL-2+IL-18	0.0	0.0	NCI-H292 IL-9	0.0	0.0
LAK cells PMA/ionomycin	0.0	0.0	NCI-H292 IL-13	0.0	0.0
NK Cells IL-2 rest	0.0	0.0	NCI-H292 IFN gamma	0.0	0.0
Two Way MLR 3 day	0.0	0.0	HPAEC none	0.0	0.0
Two Way MLR 5 day	0.0	0.0	HPAEC TNF alpha + IL-1 beta	0.0	0.0
Two Way MLR 7 day	0.0	0.0	Lung fibroblast none	0.0	0.0
PBMC rest	0.0	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0	0.0
PBMC PWM	0.0	0.0	Lung fibroblast IL-4	1.1	0.0
PBMC PHA-L	0.0	0.0	Lung fibroblast IL-9	0.0	0.0
Ramos (B cell) none	0.0	0.0	Lung fibroblast IL-13	2.5	0.0
Ramos (B cell) ionomycin	0.0	0.0	Lung fibroblast IFN gamma	0.0	0.0

B lymphocytes PWM	0.0	0.0	Dermal fibroblast CCD1070 rest	0.0	0.0
B lymphocytes CD40L and IL-4	0.0	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0	0.0
EOL-1 dbcAMP	0.0	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	0.0	Dermal fibroblast IFN gamma	1.4	0.0
Dendritic cells none	0.0	0.0	Dermal fibroblast IL-4	2.1	0.0
Dendritic cells LPS	0.0	0.0	Dermal Fibroblasts rest	0.0	0.0
Dendritic cells anti-CD40	0.0	0.0	Neutrophils TNFa+LPS	2.8	0.0
Monocytes rest	0.0	0.0	Neutrophils rest	0.0	0.0
Monocytes LPS	0.0	0.0	Colon	1.0	30.4
Macrophages rest	0.0	0.0	Lung	0.0	0.0
Macrophages LPS	0.0	0.0	Thymus	10.8	0.0
HUVEC none	0.0	0.0	Kidney	100.0	100.0
HUVEC starved	0.0	0.0			

CNS_neurodegeneration_v1.0 Summary: Ag4197/Ag4299 Expression of the CG100872-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4197 Expression of the CG100872-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4197/Ag4299 Low expression of the CG100872-01 gene is seen in kidney (CTs=32-36). Therefore, expression of this gene can be used to distinguish this sample from other samples used in this panel. In addition, therapeutic modulation of this gene can be useful in the treatment of autoimmune and inflammatory diseases that affect the kidney including lupus and glomerulonephritis.

General oncology screening_panel_v_2.4 Summary: Ag4197 Expression of the CG100872-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

N. CG100980-01: Protein-Arginine Deiminase Type III

Expression of gene CG100980-01 was assessed using the primer-probe set Ag4200, described in Table NA. Results of the RTQ-PCR runs are shown in Tables NB and NC.

Table NA. Probe Name Ag4200

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-actgcaccttcattgatgactt-3'	22	1881	245
Probe	TET-5'-actccataccacatgctgcatggg-3'-TAMRA	24	1904	246
Reverse	5'-cacttgaaagagaagggctttc-3'	22	1956	247

5

Table NB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4200, Run 221178493	Tissue Name	Rel. Exp.(%) Ag4200, Run 221178493
Adipose	0.0	Renal ca. TK-10	0.1
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.8
Melanoma* M14	0.0	Gastric ca. KATO III	0.1
Melanoma* LOXIMV1	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	2.1
Squamous cell carcinoma SCC-4	2.4	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.1	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.1	Colon ca. HCT-116	0.1
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.4	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.1	Colon Pool	0.0
Ovarian ca. OVCAR-5	3.6	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.2	Stomach Pool	0.0

Ovarian ca. OVCA-8	8.5	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.9	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.2	Fetal Skeletal Muscle	0.0
Breast ca. T47D	8.1	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	0.0
Trachea	0.0	CNS cancer (glio/astro) U87-MG	6.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.0	CNS cancer (neuro:met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.2	CNS cancer (astro) SNB-75	0.4
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.3
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	2.4
Lung ca. A549	0.5	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.1	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.2	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.0	Brain (whole)	0.0
Liver ca. HepG2	0.0	Spinal Cord Pool	0.0
Kidney Pool	0.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.6	Thyroid (female)	0.0
Renal ca. ACHN	100.0	Pancreatic ca. CAPAN2	0.1
Renal ca. UO-31	0.2	Pancreas Pool	0.0

Table NC. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4200, Run 268695263	Tissue Name	Rel. Exp.(%) Ag4200, Run 268695263
-------------	--	-------------	--

Colon cancer 1	0.2	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	6.2
Colon cancer 2	2.1	Bladder cancer NAT 4	0.2
Colon cancer NAT 2	0.2	Adenocarcinoma of the prostate 1	4.9
Colon cancer 3	0.2	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.2	Adenocarcinoma of the prostate 3	0.2
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	11.7
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.8
Lung cancer 1	0.1	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	0.2
Lung cancer 2	0.4	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	7.9	Prostate cancer NAT 10	1.0
Lung NAT 3	0.0	Kidney cancer 1	0.2
metastatic melanoma 1	0.0	Kidney NAT 1	0.0
Melanoma 2	0.9	Kidney cancer 2	0.6
Melanoma 3	0.4	Kidney NAT 2	0.0
metastatic melanoma 4	0.2	Kidney cancer 3	0.0
metastatic melanoma 5	0.0	Kidney NAT 3	0.0
Bladder cancer 1	0.2	Kidney cancer 4	0.2
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	100.0		

CNS_neurodegeneration_v1.0 Summary: Ag4200 Expression of the CG100980-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4200 Highest expression of the CG100980-01 gene is detected in renal cancer ACHN cell line (CT=23.6). Significant

expression of this gene is seen exclusively in cluster of CNS, pancreatic, colon, renal, lung, breast, ovarian, squamous cell carcinoma and prostate cancers cell lines. Therefore, expression of this gene can be used as diagnostic marker for these cancers and therapeutic modulation through the use of small molecule target could be useful in the treatments of these cancers.

Panel 4.1D Summary: Ag4200 Expression of the CG100980-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 5 Islet Summary: Ag4200 Expression of the CG100980-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4200 Highest expression of the CG100980-01 gene is detected in bladder cancer (CT=28.7), with low but significant expression of this gene in squamous cell carcinoma, colon and prostate cancers. In addition, expression of this gene is higher in the cancers than in the normal adjacent tissue. Therefore, expression of this gene may be useful as a marker to detect the presence of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of these cancers.

O. CG56763-01: GPCR

Expression of gene CG56763-01 was assessed using the primer-probe set Ag3012, described in Table OA. Results of the RTQ-PCR runs are shown in Tables OB and OC.

Table OA. Probe Name Ag3012

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctctttgtcctggtggagaac-3'	21	162	248
Probe	TET-5'-acctccctccacaggcccatgtacta-3'-TAMRA	26	213	249
Reverse	5'-gaaagacatggagctcagaaag-3'	22	239	250

Table OB. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag3012, Run 167810404	Tissue Name	Rel. Exp.(%) Ag3012, Run 167810404
Liver adenocarcinoma	0.0	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	1.6	Renal ca. A498	0.0
Adrenal gland	0.0	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0

Salivary gland	2.8	Renal ca. UO-31	0.0
Pituitary gland	0.0	Renal ca. TK-10	0.0
Brain (fetal)	0.0	Liver	0.0
Brain (whole)	0.0	Liver (fetal)	0.0
Brain (amygdala)	12.6	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.0	Lung	0.0
Brain (hippocampus)	5.5	Lung (fetal)	0.0
Brain (substantia nigra)	3.0	Lung ca. (small cell) LX-1	1.4
Brain (thalamus)	3.1	Lung ca. (small cell) NCI- H69	0.0
Cerebral Cortex	0.0	Lung ca. (s.cell var.) SHP- 77	2.4
Spinal cord	100.0	Lung ca. (large cell) NCI- H460	0.0
glio/astro U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
glio/astro U-118-MG	0.0	Lung ca. (non-s.cell) NCI- H23	0.0
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP- 62	0.0
neuro*; met SK-N-AS	0.0	Lung ca. (non-s.cl) NCI- H522	0.0
astrocytoma SF-539	0.0	Lung ca. (squam.) SW 900	0.0
astrocytoma SNB-75	0.0	Lung ca. (squam.) NCI- H596	0.0
glioma SNB-19	3.2	Mammary gland	2.9
glioma U251	0.0	Breast ca.* (pl.ef) MCF-7	3.5
glioma SF-295	0.0	Breast ca.* (pl.ef) MDA- MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl.ef) T47D	0.0
Heart	2.6	Breast ca. BT-549	0.0
Skeletal muscle (fetal)	12.5	Breast ca. MDA-N	0.0
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	3.0
Thymus	9.5	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	3.5
Lymph node	0.0	Ovarian ca. OVCAR-8	0.0
Colorectal	4.3	Ovarian ca. IGROV-1	0.0
Stomach	0.0	Ovarian ca.* (ascites) SK- OV-3	0.0

Small intestine	0.0	Uterus	0.0
Colon ca. SW480	0.0	Placenta	0.0
Colon ca.* SW620(SW480 met)	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	35.4
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	6.0	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	0.0	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK- MEL-5	0.0
Kidney	0.0	Adipose	0.0

Table OC. Panel 4D

Tissue Name	Rel. Exp.(%) Ag3012, Run 164404080	Tissue Name	Rel. Exp.(%) Ag3012, Run 164404080
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	8.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0

CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	7.2
Secondary CD8 lymphocyte rest	7.9	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	10.2	KU-812 (Basophil) PMA/ionomycin	7.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	10.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	0.0	Lupus kidney	0.0
LAK cells IL-2+IFN gamma	9.4	NCI-H292 none	0.0
LAK cells IL-2+ IL-18	14.1	NCI-H292 IL-4	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-9	18.9
NK Cells IL-2 rest	0.0	NCI-H292 IL-13	0.0
Two Way MLR 3 day	4.0	NCI-H292 IFN gamma	0.0
Two Way MLR 5 day	0.0	HPAEC none	0.0
Two Way MLR 7 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	0.0	Lung fibroblast none	0.0
PBMC PWM	8.1	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	0.0
B lymphocytes PWM	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes CD40L and IL-4	19.6	Dermal fibroblast CCD1070 rest	0.0
EOL-1 dbcAMP	16.7	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
Dendritic cells none	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells LPS	4.5	Dermal fibroblast IL-4	0.0
Dendritic cells anti-	0.0	IBD Colitis 2	0.0

CD40			
Monocytes rest	0.0	IBD Crohn's	57.0
Monocytes LPS	0.0	Colon	56.6
Macrophages rest	0.0	Lung	65.5
Macrophages LPS	0.0	Thymus	14.9
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag3012 Expression of this gene is low/undetectable (CTs > 34.5) across all of the samples on this panel (data not shown).

Panel 1.3D Summary: Ag3012 This gene is expressed at low levels in the samples derived from spinal cord (CT = 32.1) and testis (CT=33.6). Thus, the expression of this gene could be used to distinguish these samples from the other samples in the panel. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of CNS disorders, fertility and hypogonadism.

Panel 4D Summary: Ag3012 This gene is expressed at low levels in a sample derived from liver cirrhosis (CT = 33.4). In addition, expression of this gene is not detected in normal liver in Panel 1.3D, suggesting that its expression is unique to liver cirrhosis. This gene encodes a putative GPCR; therefore, antibodies or small molecule therapeutics could reduce or inhibit fibrosis that occurs in liver cirrhosis. This gene is also expressed at low levels in the lung and colon.

General oncology screening panel_v_2.4 Summary: Ag3012 Expression of this gene is low/undetectable (CTs > 34.5) across all of the samples on this panel (data not shown).

P. CG56777-01: Prostaglandin-F Synthase 1

Expression of gene CG56777-01 was assessed using the primer-probe set Ag3017, described in Table PA.

Table PA. Probe Name Ag3017

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agaaacttgaccggactatgt-3'	22	349	251
Probe	TET-5'-tcatgtaccatttgcctatgaagcctg-3'-TAMRA	26	386	252
Reverse	5'-tcctttggcagtaattctttcc-3'	22	412	253

CNS_neurodegeneration_v1.0 Summary: Ag3017 Results from one experiment with the CG56777-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 1.3D Summary: Ag3017 Expression of the CG56777-01 gene is
5 low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4D Summary: Ag3017 Expression of the CG56777-01 gene is
low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag3017 Expression of the CG56777-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel
10 (data not shown).

Q. CG56941-01: >ptnr:SPTREMBL-ACC:O60523 Ribonuclease H Type II

Expression of gene CG56941-01 was assessed using the primer-probe set Ag3096.
described in Table QA. Results of the RTQ-PCR runs are shown in Tables QB, QC, QD
15 and QE.

Table QA. Probe Name Ag3096

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtttcagaagggcaggaaaa-3'	20	321	254
Probe	TET-5'-caacatggacaagaatcggagacgaa-3'-TAMRA	26	342	255
Reverse	5'-ttcatctccatctccatcca-3'	20	394	256

Table QB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3096, Run 208976837	Tissue Name	Rel. Exp.(%) Ag3096, Run 208976837
AD 1 Hippo	18.9	Control (Path) 3 Temporal Ctx	11.3
AD 2 Hippo	17.1	Control (Path) 4 Temporal Ctx	43.2
AD 3 Hippo	7.1	AD 1 Occipital Ctx	46.7
AD 4 Hippo	5.1	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	100.0	AD 3 Occipital Ctx	6.7
AD 6 Hippo	34.6	AD 4 Occipital Ctx	29.3
Control 2 Hippo	13.8	AD 5 Occipital Ctx	20.3

Control 4 Hippo	8.0	AD 6 Occipital Ctx	27.2
Control (Path) 3 Hippo	6.0	Control 1 Occipital Ctx	11.5
AD 1 Temporal Ctx	16.2	Control 2 Occipital Ctx	4.1
AD 2 Temporal Ctx	25.5	Control 3 Occipital Ctx	49.0
AD 3 Temporal Ctx	3.4	Control 4 Occipital Ctx	4.0
AD 4 Temporal Ctx	22.5	Control (Path) 1 Occipital Ctx	61.6
AD 5 Inf Temporal Ctx	55.5	Control (Path) 2 Occipital Ctx	18.2
AD 5 Sup Temporal Ctx	33.7	Control (Path) 3 Occipital Ctx	5.5
AD 6 Inf Temporal Ctx	45.1	Control (Path) 4 Occipital Ctx	79.6
AD 6 Sup Temporal Ctx	61.1	Control 1 Parietal Ctx	9.6
Control 1 Temporal Ctx	8.9	Control 2 Parietal Ctx	58.2
Control 2 Temporal Ctx	20.4	Control 3 Parietal Ctx	15.9
Control 3 Temporal Ctx	18.8	Control (Path) 1 Parietal Ctx	37.9
Control 3 Temporal Ctx	12.9	Control (Path) 2 Parietal Ctx	18.3
Control (Path) 1 Temporal Ctx	46.0	Control (Path) 3 Parietal Ctx	2.2
Control (Path) 2 Temporal Ctx	18.6	Control (Path) 4 Parietal Ctx	59.0

Table QC. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag3096, Run 167985249	Tissue Name	Rel. Exp.(%) Ag3096, Run 167985249
Liver adenocarcinoma	0.5	Kidney (fetal)	0.2
Pancreas	0.0	Renal ca. 786-0	0.2
Pancreatic ca. CAPAN 2	0.1	Renal ca. A498	0.2
Adrenal gland	0.0	Renal ca. RXF 393	0.1
Thyroid	0.0	Renal ca. ACHN	0.1
Salivary gland	0.1	Renal ca. UO-31	0.2
Pituitary gland	0.5	Renal ca. TK-10	0.1
Brain (fetal)	83.5	Liver	0.0

Brain (whole)	0.7	Liver (fetal)	0.0
Brain (amygdala)	1.1	Liver ca. (hepatoblast) HepG2	3.7
Brain (cerebellum)	100.0	Lung	0.0
Brain (hippocampus)	1.1	Lung (fetal)	0.1
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.3
Brain (thalamus)	0.3	Lung ca. (small cell) NCI-H69	0.1
Cerebral Cortex	1.0	Lung ca. (s.cell var.) SHP-77	1.4
Spinal cord	0.1	Lung ca. (large cell) NCI-H460	0.4
glio/astro U87-MG	0.5	Lung ca. (non-sm. cell) A549	1.2
glio/astro U-118-MG	0.1	Lung ca. (non-s.cell) NCI-H23	0.3
astrocytoma SW1783	0.2	Lung ca. (non-s.cell) HOP-62	0.1
neuro*; met SK-N-AS	4.1	Lung ca. (non-s.cl) NCI-H522	0.3
astrocytoma SF-539	0.1	Lung ca. (squam.) SW 900	0.1
astrocytoma SNB-75	0.4	Lung ca. (squam.) NCI-H596	0.3
glioma SNB-19	0.1	Mammary gland	0.0
glioma U251	0.4	Breast ca.* (pl.ef) MCF-7	0.3
glioma SF-295	0.3	Breast ca.* (pl.ef) MDA-MB-231	0.1
Heart (fetal)	0.0	Breast ca.* (pl.ef) T47D	0.1
Heart	0.1	Breast ca. BT-549	0.2
Skeletal muscle (fetal)	0.0	Breast ca. MDA-N	0.2
Skeletal muscle	0.1	Ovary	0.0
Bone marrow	0.0	Ovarian ca. OVCAR-3	4.4
Thymus	0.0	Ovarian ca. OVCAR-4	0.1
Spleen	0.0	Ovarian ca. OVCAR-5	0.4
Lymph node	0.1	Ovarian ca. OVCAR-8	0.5
Colorectal	0.0	Ovarian ca. IGROV-1	0.5
Stomach	0.1	Ovarian ca.* (ascites) SK-OV-3	0.6
Small intestine	0.0	Uterus	0.1

Colon ca. SW480	0.4	Placenta	0.0
Colon ca.* SW620(SW480 met)	0.8	Prostate	0.0
Colon ca. HT29	0.1	Prostate ca.* (bone met)PC-3	0.3
Colon ca. HCT-116	0.3	Testis	1.2
Colon ca. CaCo-2	0.5	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	0.1	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC-2998	0.3	Melanoma UACC-62	0.1
Gastric ca.* (liver met) NCI-N87	0.6	Melanoma M14	0.1
Bladder	2.0	Melanoma LOX IMVI	0.3
Trachea	0.0	Melanoma* (met) SK-MEL-5	0.1
Kidney	0.1	Adipose	0.1

Table QD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3096. Run 169990855	Tissue Name	Rel. Exp.(%) Ag3096. Run 169990855
Secondary Th1 act	41.2	HUVEC IL-1beta	45.4
Secondary Th2 act	70.7	HUVEC IFN gamma	35.1
Secondary Tr1 act	65.5	HUVEC TNF alpha + IFN gamma	38.7
Secondary Th1 rest	14.4	HUVEC TNF alpha + IL4	33.4
Secondary Th2 rest	46.0	HUVEC IL-11	18.4
Secondary Tr1 rest	34.2	Lung Microvascular EC none	46.0
Primary Th1 act	54.7	Lung Microvascular EC TNFalpha + IL-1beta	39.5
Primary Th2 act	57.0	Microvascular Dermal EC none	100.0
Primary Tr1 act	44.1	Microvascular Dermal EC TNFalpha + IL-1beta	40.9
Primary Th1 rest	12.0	Bronchial epithelium TNFalpha + IL1beta	17.3
Primary Th2 rest	18.2	Small airway epithelium none	21.6
Primary Tr1 rest	17.8	Small airway epithelium TNFalpha + IL-1beta	34.4

CD45RA CD4 lymphocyte act	67.4	Coronary artery SMC rest	37.1
CD45RO CD4 lymphocyte act	60.3	Coronary artery SMC TNFalpha + IL-1beta	27.5
CD8 lymphocyte act	55.1	Astrocytes rest	20.2
Secondary CD8 lymphocyte rest	46.7	Astrocytes TNFalpha + IL-1beta	26.1
Secondary CD8 lymphocyte act	42.0	KU-812 (Basophil) rest	18.2
CD4 lymphocyte none	26.4	KU-812 (Basophil) PMA/ionomycin	31.2
2ry Th1/Th2/Tr1_anti-CD95 CH11	17.9	CCD1106 (Keratinocytes) none	62.4
LAK cells rest	31.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	25.2
LAK cells IL-2	41.5	Liver cirrhosis	11.8
LAK cells IL-2+IL-12	53.6	NCI-H292 none	23.3
LAK cells IL-2+IFN gamma	59.5	NCI-H292 IL-4	29.9
LAK cells IL-2+ IL-18	36.6	NCI-H292 IL-9	45.4
LAK cells PMA/ionomycin	36.3	NCI-H292 IL-13	48.6
NK Cells IL-2 rest	42.9	NCI-H292 IFN gamma	48.0
Two Way MLR 3 day	69.3	HPAEC none	28.5
Two Way MLR 5 day	79.0	HPAEC TNF alpha + IL-1 beta	46.0
Two Way MLR 7 day	21.5	Lung fibroblast none	26.4
PBMC rest	13.7	Lung fibroblast TNF alpha + IL-1 beta	30.6
PBMC PWM	46.0	Lung fibroblast IL-4	34.4
PBMC PHA-L	18.3	Lung fibroblast IL-9	62.9
Ramos (B cell) none	59.0	Lung fibroblast IL-13	62.0
Ramos (B cell) ionomycin	49.3	Lung fibroblast IFN gamma	28.5
B lymphocytes PWM	54.0	Dermal fibroblast CCD1070 rest	58.6
B lymphocytes CD40L and IL-4	37.6	Dermal fibroblast CCD1070 TNF alpha	57.4

EOL-1 dbcAMP	22.8	Dermal fibroblast CCD1070 IL-1 beta	33.7
EOL-1 dbcAMP PMA/ionomycin	20.2	Dermal fibroblast IFN gamma	12.9
Dendritic cells none	81.8	Dermal fibroblast IL-4	35.6
Dendritic cells LPS	15.5	Dermal Fibroblasts rest	22.8
Dendritic cells anti- CD40	14.9	Neutrophils TNFa+LPS	0.0
Monocytes rest	25.9	Neutrophils rest	3.4
Monocytes LPS	37.1	Colon	3.5
Macrophages rest	36.9	Lung	11.2
Macrophages LPS	14.8	Thymus	16.8
HUVEC none	47.6	Kidney	14.6
HUVEC starved	28.3		

Table QE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3096, Run 267920136	Tissue Name	Rel. Exp.(%) Ag3096, Run 267920136
Colon cancer 1	11.8	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	4.8	Bladder cancer NAT 3	0.0
Colon cancer 2	7.2	Bladder cancer NAT 4	3.3
Colon cancer NAT 2	7.8	Adenocarcinoma of the prostate 1	31.2
Colon cancer 3	22.7	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	12.2	Adenocarcinoma of the prostate 3	3.2
Colon malignant cancer 4	6.4	Adenocarcinoma of the prostate 4	12.0
Colon normal adjacent tissue 4	5.0	Prostate cancer NAT 5	0.0
Lung cancer 1	10.8	Adenocarcinoma of the prostate 6	2.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	9.0
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	2.6	Adenocarcinoma of the prostate 9	11.4
Squamous cell carcinoma 3	17.2	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	9.3

metastatic melanoma 1	4.6	Kidney NAT 1	10.5
Melanoma 2	0.0	Kidney cancer 2	40.6
Melanoma 3	2.6	Kidney NAT 2	7.7
metastatic melanoma 4	12.8	Kidney cancer 3	5.4
metastatic melanoma 5	20.0	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	10.4
Bladder cancer NAT 1	0.0	Kidney NAT 4	5.4
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag3096 This panel confirms the expression of the CG56941-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

Panel 1.3D Summary: Ag3096 Highest expression of the CG56941-01 gene is detected in brain (cerebellum) (CT=27.7). In addition, high to moderate expression of this gene is also detected in fetal brain and other regions of central nervous systems, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Low levels of expression of this gene is also seen in a ovarian cancer, two lung cancer, liver cancer, a colon cancer and a CNS cancer cell line. Therefore, therapeutic modulation of this gene could be useful in the treatment of these cancers.

Panel 4.1D Summary: Ag3096 Highest expression of the CG56941-01 gene is detected in microvascular dermal endothelial cells (CT=34.4). In addition, low levels of expression of this gene is also seen in dendritic cells, and two way MLR. Therefore, therapeutic modulation of this gene could be useful in the treatment of autoimmune and inflammatory diseases that involve these cells, such as lupus erythematosus, asthma, emphysema, Crohn's disease, ulcerative colitis, rheumatoid arthritis, osteoarthritis, and psoriasis.

General oncology screening panel_v_2.4 Summary: Ag3096 Highest expression of the CG56941-01 gene is detected in lung cancer (CT=34.3). Expression of this gene is higher in the cancer sample than in the corresponding adjacent control sample (CT=39.6). Thus, expression of this gene may be useful as diagnostic marker for detection of lung cancer and therapeutic modulation of this gene may be useful in the treatment of this cancer.

R. CG57109-01 and CG57109-02 and CG57109-03 and CG57109-04 and CG57109-05 and CG57109-06: Doublecortin/CAMKinase

Expression of gene CG57109-01, variants CG57109-02, CG57109-03, CG57109-04, CG57109-06 and full length physical clone CG57109-05 was assessed using the primer-probe sets Ag1137, Ag1150, Ag1860, Ag3112 and Ag4281, described in Tables RA, RB, RC, RD and RE. Results of the RTQ-PCR runs are shown in Tables RF, RG, RH, RI, RJ, RK, RL, RM and RN. Please note that the variants CG57109-03, CG57109-04, CG57109-06 correspond to the probe and primer sets Ag150 and Ag3112 only. CG57109-05 represents a full-length physical clone of the CG57109-01 gene, validating the prediction of the gene sequence.

Table RA. Probe Name Ag1137

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gacatggtggacagtgagatct-3'	22	1338	257
Probe	TET-5'-cctctctcaccccaacatcgtagaaat-3'-TAMRA	26	1373	258
Reverse	5'-tctgtttcgttagacttcatgca-3'	22	1399	259

Table RB. Probe Name Ag1150

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gaaattggctgattttggactt-3'	22	1634	260
Probe	TET-5'-cctatatattactgtgtgtgtgggacccca-3'-TAMRA	27	1674	261
Reverse	5'-agaatttcgggagctacgtaag-3'	22	1702	262

Table RC. Probe Name Ag1860

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gacatggtggacagtgagatct-3'	22	1338	263
Probe	TET-5'-cctctctcaccccaacatcgtagaaat-3'-TAMRA	26	1373	264
Reverse	5'-tctgtttcgttagacttcatgca-3'	22	1399	265

Table RD. Probe Name Ag3112

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gaaattggctgattttggactt-3'	22	1634	266
Probe	TET-5'-cctatatttactgtgtgtgggacccca-3'-TAMRA	27	1674	267
Reverse	5'-agaatttcgggagctacgtaag-3'	22	1702	268

Table RE. Probe Name Ag4281

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gacatgggtggacagtgagatct-3'	22	1338	269
Probe	TET-5'-atccagagcctctctcaccaccaacat-3'-TAMRA	26	1365	270
Reverse	5'-tcgtagacttcatgcaatttca-3'	22	1393	271

Table RF. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag1860, Run 225404259	Tissue Name	Rel. Exp.(%) Ag1860, Run 225404259
110967 COPD-F	3.7	112427 Match Control Psoriasis-F	7.9
110980 COPD-F	0.0	112418 Psoriasis-M	13.1
110968 COPD-M	20.0	112723 Match Control Psoriasis-M	17.9
110977 COPD-M	3.0	112419 Psoriasis-M	26.8
110989 Emphysema-F	26.2	112424 Match Control Psoriasis-M	19.5
110992 Emphysema-F	4.2	112420 Psoriasis-M	36.9
110993 Emphysema-F	11.0	112425 Match Control Psoriasis-M	23.3
110994 Emphysema-F	19.8	104689 (MF) OA Bone- Backus	13.0
110995 Emphysema-F	8.5	104690 (MF) Adj "Normal" Bone-Backus	3.9
110996 Emphysema-F	0.0	104691 (MF) OA Synovium-Backus	29.5
110997 Asthma-M	17.2	104692 (BA) OA Cartilage-Backus	0.0
111001 Asthma-F	38.4	104694 (BA) OA Bone- Backus	17.3
111002 Asthma-F	27.9	104695 (BA) Adj "Normal" Bone-Backus	0.0

111003 Atopic Asthma-F	50.7	104696 (BA) OA Synovium-Backus	93.3
111004 Atopic Asthma-F	68.3	104700 (SS) OA Bone-Backus	11.7
111005 Atopic Asthma-F	56.6	104701 (SS) Adj "Normal" Bone-Backus	15.3
111006 Atopic Asthma-F	17.9	104702 (SS) OA Synovium-Backus	100.0
111417 Allergy-M	31.0	117093 OA Cartilage Rep7	19.1
112347 Allergy-M	8.3	112672 OA Bone5	6.8
112349 Normal Lung-F	2.5	112673 OA Synovium5	4.8
112357 Normal Lung-F	4.4	112674 OA Synovial Fluid cells5	13.9
112354 Normal Lung-M	0.0	117100 OA Cartilage Rep14	4.5
112374 Crohns-F	27.5	112756 OA Bone9	63.7
112389 Match Control Crohns-F	9.0	112757 OA Synovium9	17.7
112375 Crohns-F	36.6	112758 OA Synovial Fluid Cells9	11.7
112732 Match Control Crohns-F	0.0	117125 RA Cartilage Rep2	27.5
112725 Crohns-M	4.4	113492 Bone2 RA	20.0
112387 Match Control Crohns-M	17.1	113493 Synovium2 RA	6.8
112378 Crohns-M	4.2	113494 Syn Fluid Cells RA	8.8
112390 Match Control Crohns-M	49.3	113499 Cartilage4 RA	4.7
112726 Crohns-M	30.1	113500 Bone4 RA	8.8
112731 Match Control Crohns-M	0.0	113501 Synovium4 RA	3.5
112380 Ulcer Col-F	37.9	113502 Syn Fluid Cells4 RA	4.0
112734 Match Control Ulcer Col-F	17.9	113495 Cartilage3 RA	10.4
112384 Ulcer Col-F	29.7	113496 Bone3 RA	4.2
112737 Match Control Ulcer Col-F	9.0	113497 Synovium3 RA	0.0
112386 Ulcer Col-F	5.0	113498 Syn Fluid Cells3 RA	7.9

112738 Match Control Ulcer Col-F	9.8	117106 Normal Cartilage Rep20	0.0
112381 Ulcer Col-M	0.0	113663 Bone3 Normal	0.0
112735 Match Control Ulcer Col-M	28.7	113664 Synovium3 Normal	0.0
112382 Ulcer Col-M	10.0	113665 Syn Fluid Cells3 Normal	0.0
112394 Match Control Ulcer Col-M	0.0	117107 Normal Cartilage Rep22	0.0
112383 Ulcer Col-M	35.4	113667 Bone4 Normal	13.3
112736 Match Control Ulcer Col-M	6.3	113668 Synovium4 Normal	0.0
112423 Psoriasis-F	21.5	113669 Syn Fluid Cells4 Normal	9.4

Table RG. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag1860, Run 207807786	Rel. Exp.(%) Ag3112, Run 208976863	Rel. Exp.(%) Ag4281, Run 224075294	Tissue Name	Rel. Exp.(%) Ag1860 , Run 207807 786	Rel. Exp.(%) Ag3112, Run 208976863	Rel. Exp.(%) Ag4281, Run 224075294
AD 1 Hippo	7.6	2.7	9.7	Control (Path) 3 Temporal Ctx	0.5	2.9	3.1
AD 2 Hippo	29.1	21.8	24.5	Control (Path) 4 Temporal Ctx	28.7	23.0	27.9
AD 3 Hippo	1.5	3.2	7.1	AD 1 Occipital Ctx	17.6	9.8	27.4
AD 4 Hippo	4.2	1.1	5.6	AD 2 Occipital Ctx (Missing)	0.0	0.0	0.0
AD 5 hippo	100.0	100.0	95.3	AD 3 Occipital Ctx	4.1	3.0	1.3
AD 6 Hippo	19.9	12.2	26.1	AD 4 Occipital Ctx	7.9	4.8	8.5
Control 2 Hippo	27.2	15.9	39.0	AD 5 Occipital Ctx	13.8	44.1	50.0

Control 4 Hippo	3.1	2.1	7.2	AD 6 Occipital Ctx	33.4	8.3	14.8
Control (Path) 3 Hippo	2.1	0.0	5.0	Control 1 Occipital Ctx	0.0	0.0	2.0
AD 1 Temporal Ctx	2.9	4.1	9.3	Control 2 Occipital Ctx	62.0	81.2	100.0
AD 2 Temporal Ctx	21.6	27.0	23.5	Control 3 Occipital Ctx	5.0	2.2	15.5
AD 3 Temporal Ctx	2.7	0.0	6.3	Control 4 Occipital Ctx	0.5	0.0	2.4
AD 4 Temporal Ctx	14.6	15.6	12.8	Control (Path) 1 Occipital Ctx	34.6	42.6	43.8
AD 5 Inf Temporal Ctx	38.2	48.3	54.7	Control (Path) 2 Occipital Ctx	6.2	7.1	6.7
AD 5 Sup Temporal Ctx	26.8	26.2	27.0	Control (Path) 3 Occipital Ctx	0.0	0.0	0.6
AD 6 Inf Temporal Ctx	12.6	11.7	25.2	Control (Path) 4 Occipital Ctx	15.8	4.7	25.7
AD 6 Sup Temporal Ctx	17.7	10.9	17.4	Control 1 Parietal Ctx	3.7	1.0	1.2
Control 1 Temporal Ctx	4.0	1.1	4.6	Control 2 Parietal Ctx	18.8	12.9	19.9
Control 2 Temporal Ctx	22.5	12.6	23.0	Control 3 Parietal Ctx	8.1	2.7	7.9
Control 3 Temporal Ctx	7.0	4.8	4.8	Control (Path) 1 Parietal Ctx	33.4	44.1	42.9
Control 4 Temporal Ctx	4.8	1.9	6.5	Control (Path) 2 Parietal	17.4	11.0	13.6

				Ctx			
Control (Path) 1 Temporal Ctx	37.6	34.9	47.0	Control (Path) 3 Parietal Ctx	1.7	0.0	0.7
Control (Path) 2 Temporal Ctx	28.1	13.6	27.9	Control (Path) 4 Parietal Ctx	19.9	6.9	24.8

Table RH. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4281, Run 222183233	Tissue Name	Rel. Exp.(%) Ag4281, Run 222183233
Adipose	0.5	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	5.4
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK- MEL-5	0.0	Colon ca. SW480	0.2
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	34.2	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.9	Colon ca. CaCo-2	2.6
Placenta	2.0	Colon cancer tissue	5.6
Uterus Pool	0.8	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.7	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV- 3	0.6	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	4.0
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.5
Ovarian ca. IGROV- 1	0.0	Stomach Pool	0.4
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	1.1

Ovary	0.6	Fetal Heart	4.5
Breast ca. MCF-7	0.0	Heart Pool	2.7
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	5.9
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	22.1
Breast ca. T47D	1.0	Skeletal Muscle Pool	9.3
Breast ca. MDA-N	0.0	Spleen Pool	2.4
Breast Pool	3.1	Thymus Pool	3.1
Trachea	1.7	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	14.4	CNS cancer (neuro;met) SK-N-AS	0.8
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.9	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	25.5
Lung ca. NCI-H526	0.0	Brain (cerebellum)	5.4
Lung ca. NCI-H23	2.1	Brain (fetal)	100.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	59.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	14.1
Lung ca. NCI-H522	0.7	Brain (Substantia nigra) Pool	15.1
Liver	0.3	Brain (Thalamus) Pool	28.9
Fetal Liver	1.0	Brain (whole)	33.9
Liver ca. HepG2	0.0	Spinal Cord Pool	25.2
Kidney Pool	4.2	Adrenal Gland	2.3
Fetal Kidney	4.6	Pituitary gland Pool	36.6
Renal ca. 786-0	0.0	Salivary Gland	0.4
Renal ca. A498	0.0	Thyroid (female)	0.4
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	2.1

Table RI. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag1860, Run 165981809	Rel. Exp.(%) Ag3112, Run 167985258	Tissue Name	Rel. Exp.(%) Ag1860, Run 165981809	Rel. Exp.(%) Ag3112, Run
-------------	--	--	-------------	--	--------------------------------

					167985258
Liver adenocarcinoma	0.0	0.0	Kidney (fetal)	0.8	1.4
Pancreas	0.0	0.0	Renal ca. 786-0	0.0	0.0
Pancreatic ca. CAPAN 2	0.0	0.0	Renal ca. A498	0.0	0.0
Adrenal gland	0.0	0.0	Renal ca. RXF 393	0.0	0.0
Thyroid	0.0	0.0	Renal ca. ACHN	0.0	0.0
Salivary gland	0.0	0.0	Renal ca. UO-31	0.0	0.0
Pituitary gland	12.6	6.7	Renal ca. TK-10	0.0	0.0
Brain (fetal)	26.6	99.3	Liver	0.0	0.0
Brain (whole)	7.0	8.3	Liver (fetal)	0.0	0.0
Brain (amygdala)	35.6	19.3	Liver ca. (hepatoblast) HepG2	0.0	0.0
Brain (cerebellum)	1.2	1.3	Lung	0.0	0.0
Brain (hippocampus)	7.0	19.5	Lung (fetal)	1.0	2.4
Brain (substantia nigra)	0.4	6.7	Lung ca. (small cell) LX-1	0.0	0.0
Brain (thalamus)	5.8	0.6	Lung ca. (small cell) NCI-H69	0.0	0.0
Cerebral Cortex	13.6	9.3	Lung ca. (s.cell var.) SHP-77	0.0	0.0
Spinal cord	100.0	100.0	Lung ca. (large cell) NCI-H460	0.0	0.0
glio/astro U87-MG	0.0	0.0	Lung ca. (non-sm. cell) A549	0.0	1.3
glio/astro U-118-MG	0.0	0.0	Lung ca. (non-s.cell) NCI-H23	2.1	0.0
astrocytoma SW1783	0.0	0.0	Lung ca. (non-s.cell) HOP-62	0.0	0.0
neuro*; met SK-N-AS	0.0	0.0	Lung ca. (non-s.cl) NCI-H522	0.0	0.0
astrocytoma SF-539	0.0	0.6	Lung ca. (squam.) SW 900	0.0	0.0
astrocytoma SNB-75	0.0	0.0	Lung ca. (squam.) NCI-H596	0.0	0.0
glioma SNB-19	0.0	0.0	Mammary gland	0.0	0.0
glioma U251	0.0	0.0	Breast ca.* (pl.ef) MCF-7	0.0	0.0
glioma SF-295	0.0	0.0	Breast ca.* (pl.ef)	0.0	0.0

			MDA-MB-231		
Heart (fetal)	0.0	0.0	Breast ca.* (pl.ef) T47D	0.0	0.0
Heart	0.7	4.5	Breast ca. BT-549	0.0	0.0
Skeletal muscle (fetal)	4.6	19.2	Breast ca. MDA-N	0.0	0.0
Skeletal muscle	4.9	8.1	Ovary	0.0	0.0
Bone marrow	0.0	0.0	Ovarian ca. OVCAR-3	0.0	0.0
Thymus	1.1	2.2	Ovarian ca. OVCAR-4	0.0	0.0
Spleen	0.0	0.0	Ovarian ca. OVCAR-5	0.0	0.0
Lymph node	0.7	1.3	Ovarian ca. OVCAR-8	0.0	0.0
Colorectal	2.8	0.3	Ovarian ca. IGROV-1	0.0	0.0
Stomach	0.0	1.1	Ovarian ca.* (ascites) SK-OV-3	0.0	1.7
Small intestine	1.1	2.0	Uterus	0.6	0.0
Colon ca. SW480	0.0	0.0	Placenta	2.8	0.0
Colon ca.* SW620(SW480 met)	0.0	0.0	Prostate	0.0	0.0
Colon ca. HT29	0.0	0.0	Prostate ca.* (bone met)PC-3	0.0	0.0
Colon ca. HCT- 116	0.0	0.0	Testis	42.6	36.1
Colon ca. CaCo-2	0.7	0.0	Melanoma Hs688(A).T	0.0	0.0
Colon ca. tissue(ODO3866)	2.0	4.0	Melanoma* (met) Hs688(B).T	0.0	0.0
Colon ca. HCC- 2998	0.0	0.0	Melanoma UACC- 62	0.0	0.0
Gastric ca.* (liver met) NCI-N87	0.0	0.0	Melanoma M14	0.0	0.0
Bladder	4.0	1.3	Melanoma LOX IMVI	0.0	0.0
Trachea	3.0	1.4	Melanoma* (met) SK-MEL-5	0.0	0.0
Kidney	0.0	0.0	Adipose	1.3	2.6

Table RJ. Panel 2.2

Tissue Name	Rel. Exp.(%)	Tissue Name	Rel. Exp.(%)
-------------	--------------	-------------	--------------

	Ag1860, Run 174229165		Ag1860, Run 174229165
Normal Colon	20.4	Kidney Margin (OD04348)	0.0
Colon cancer (OD06064)	0.0	Kidney malignant cancer (OD06204B)	28.9
Colon Margin (OD06064)	12.2	Kidney normal adjacent tissue (OD06204E)	0.0
Colon cancer (OD06159)	0.0	Kidney Cancer (OD04450- 01)	0.0
Colon Margin (OD06159)	12.6	Kidney Margin (OD04450- 03)	0.0
Colon cancer (OD06297-04)	37.6	Kidney Cancer 8120613	20.0
Colon Margin (OD06297-05)	15.0	Kidney Margin 8120614	0.0
CC Gr.2 ascend colon (ODO3921)	45.4	Kidney Cancer 9010320	17.7
CC Margin (ODO3921)	35.1	Kidney Margin 9010321	21.0
Colon cancer metastasis (OD06104)	0.0	Kidney Cancer 8120607	16.5
Lung Margin (OD06104)	19.1	Kidney Margin 8120608	0.0
Colon mets to lung (OD04451-01)	51.8	Normal Uterus	27.4
Lung Margin (OD04451-02)	0.0	Uterine Cancer 064011	0.0
Normal Prostate	0.0	Normal Thyroid	0.0
Prostate Cancer (OD04410)	0.0	Thyroid Cancer 064010	28.9
Prostate Margin (OD04410)	0.0	Thyroid Cancer A302152	33.9
Normal Ovary	0.0	Thyroid Margin A302153	0.0
Ovarian cancer (OD06283-03)	0.0	Normal Breast	28.3
Ovarian Margin (OD06283-07)	0.0	Breast Cancer (OD04566)	56.3
Ovarian Cancer 064008	30.4	Breast Cancer 1024	55.5
Ovarian cancer (OD06145)	0.0	Breast Cancer (OD04590- 01)	94.0
Ovarian Margin (OD06145)	35.1	Breast Cancer Mets (OD04590-03)	36.9
Ovarian cancer (OD06455-03)	0.0	Breast Cancer Metastasis (OD04655-05)	40.3

Ovarian Margin (OD06455-07)	0.0	Breast Cancer 064006	0.0
Normal Lung	13.3	Breast Cancer 9100266	0.0
Invasive poor diff. lung adeno (ODO4945-01)	0.0	Breast Margin 9100265	0.0
Lung Margin (ODO4945-03)	20.3	Breast Cancer A209073	37.4
Lung Malignant Cancer (OD03126)	26.2	Breast Margin A2090734	23.2
Lung Margin (OD03126)	0.0	Breast cancer (OD06083)	19.3
Lung Cancer (OD05014A)	0.0	Breast cancer node metastasis (OD06083)	60.3
Lung Margin (OD05014B)	0.0	Normal Liver	15.5
Lung cancer (OD06081)	9.4	Liver Cancer 1026	16.5
Lung Margin (OD06081)	0.0	Liver Cancer 1025	0.0
Lung Cancer (OD04237-01)	0.0	Liver Cancer 6004-T	0.0
Lung Margin (OD04237-02)	100.0	Liver Tissue 6004-N	0.0
Ocular Melanoma Metastasis	0.0	Liver Cancer 6005-T	0.0
Ocular Melanoma Margin (Liver)	0.0	Liver Tissue 6005-N	0.0
Melanoma Metastasis	0.0	Liver Cancer 064003	0.0
Melanoma Margin (Lung)	0.0	Normal Bladder	0.0
Normal Kidney	0.0	Bladder Cancer 1023	27.2
Kidney Ca, Nuclear grade 2 (OD04338)	0.0	Bladder Cancer A302173	0.0
Kidney Margin (OD04338)	0.0	Normal Stomach	17.3
Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	Gastric Cancer 9060397	20.0
Kidney Margin (OD04339)	0.0	Stomach Margin 9060396	59.0
Kidney Ca, Clear cell type (OD04340)	0.0	Gastric Cancer 9060395	14.4
Kidney Margin (OD04340)	0.0	Stomach Margin 9060394	69.7

Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer 064005	48.0
--------------------------------------	-----	-----------------------	------

Table RK. Panel 3D

Tissue Name	Rel. Exp.(%) Ag3112, Run 182114339	Tissue Name	Rel. Exp.(%) Ag3112, Run 182114339
Daoy-Medulloblastoma	0.0	Ca Ski- Cervical epidermoid carcinoma (metastasis)	0.0
TE671-Medulloblastoma	0.0	ES-2- Ovarian clear cell carcinoma	0.0
D283 Med-Medulloblastoma	0.0	Ramos- Stimulated with PMA/ionomycin 6h	0.0
PFSK-1- Primitive Neuroectodermal	0.0	Ramos- Stimulated with PMA/ionomycin 14h	0.0
XF-498- CNS	0.0	MEG-01- Chronic myelogenous leukemia (megakaryoblast)	0.0
SNB-78- Glioma	0.0	Raji- Burkitt's lymphoma	0.0
SF-268- Glioblastoma	0.0	Daudi- Burkitt's lymphoma	0.0
T98G- Glioblastoma	0.0	U266- B-cell plasmacytoma	0.0
SK-N-SH- Neuroblastoma (metastasis)	0.0	CA46- Burkitt's lymphoma	0.0
SF-295- Glioblastoma	0.0	RL- non-Hodgkin's B-cell lymphoma	0.0
Cerebellum	6.5	JM1- pre-B-cell lymphoma	0.0
Cerebellum	2.4	Jurkat- T cell leukemia	7.3
NCI-H292- Mucoepidermoid lung carcinoma	0.0	TF-1- Erythroleukemia	0.0
DMS-114- Small cell lung cancer	0.0	HUT 78- T-cell lymphoma	0.0
DMS-79- Small cell lung cancer	100.0	U937- Histiocytic lymphoma	0.0
NCI-H146- Small cell lung cancer	0.0	KU-812- Myelogenous leukemia	0.0
NCI-H526- Small cell lung cancer	0.0	769-P- Clear cell renal carcinoma	0.0
NCI-N417- Small cell lung cancer	0.0	Caki-2- Clear cell renal carcinoma	0.0
NCI-H82- Small cell lung cancer	0.0	SW 839- Clear cell renal carcinoma	0.0
NCI-H157- Squamous	0.0	G401- Wilms' tumor	0.0

cell lung cancer (metastasis)			
NCI-H1155- Large cell lung cancer	3.6	Hs766T- Pancreatic carcinoma (LN metastasis)	0.0
NCI-H1299- Large cell lung cancer	0.0	CAPAN-1- Pancreatic adenocarcinoma (liver metastasis)	0.0
NCI-H727- Lung carcinoid	0.0	SU86.86- Pancreatic carcinoma (liver metastasis)	0.0
NCI-UMC-11- Lung carcinoid	0.0	BxPC-3- Pancreatic adenocarcinoma	0.0
LX-1- Small cell lung cancer	0.0	HPAC- Pancreatic adenocarcinoma	0.0
Colo-205- Colon cancer	0.0	MIA PaCa-2- Pancreatic carcinoma	0.0
KM12- Colon cancer	0.0	CFPAC-1- Pancreatic ductal adenocarcinoma	0.0
KM20L2- Colon cancer	0.0	PANC-1- Pancreatic epithelioid ductal carcinoma	0.0
NCI-H716- Colon cancer	0.0	T24- Bladder carcinoma (transitional cell)	0.0
SW-48- Colon adenocarcinoma	0.0	5637- Bladder carcinoma	0.0
SW1116- Colon adenocarcinoma	0.0	HT-1197- Bladder carcinoma	0.0
LS 174T- Colon adenocarcinoma	0.0	UM-UC-3- Bladder carcinoma (transitional cell)	0.0
SW-948- Colon adenocarcinoma	0.0	A204- Rhabdomyosarcoma	0.0
SW-480- Colon adenocarcinoma	0.0	HT-1080- Fibrosarcoma	0.0
NCI-SNU-5- Gastric carcinoma	0.0	MG-63- Osteosarcoma	0.0
KATO III- Gastric carcinoma	0.0	SK-LMS-1- Leiomyosarcoma (vulva)	0.0
NCI-SNU-16- Gastric carcinoma	0.0	SJRH30- Rhabdomyosarcoma (met to bone marrow)	0.0
NCI-SNU-1- Gastric carcinoma	0.0	A431- Epidermoid carcinoma	0.0
RF-1- Gastric adenocarcinoma	0.0	WM266-4- Melanoma	0.0
RF-48- Gastric adenocarcinoma	0.0	DU 145- Prostate carcinoma (brain metastasis)	0.0
MKN-45- Gastric	0.0	MDA-MB-468- Breast	0.0

carcinoma		adenocarcinoma	
NCI-N87- Gastric carcinoma	0.0	SCC-4- Squamous cell carcinoma of tongue	0.0
OVCAR-5- Ovarian carcinoma	0.0	SCC-9- Squamous cell carcinoma of tongue	0.0
RL95-2- Uterine carcinoma	0.0	SCC-15- Squamous cell carcinoma of tongue	0.0
HelaS3- Cervical adenocarcinoma	0.0	CAL 27- Squamous cell carcinoma of tongue	0.0

Table RL. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4281, Run 181080824	Tissue Name	Rel. Exp.(%) Ag4281, Run 181080824
Secondary Th1 act	0.0	HUVEC IL-1beta	17.4
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	8.5
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	1.6
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	2.8	Coronary artery SMC rest	5.7
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	25.7
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	1.1
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	1.3	KU-812 (Basophil) PMA/ionomycin	0.0

2ry Th1/Th2/Tr1_anti-CD95 CH11	2.9	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	1.4
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	1.2
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	4.9
NK Cells IL-2 rest	1.8	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	1.1	HPAEC none	2.9
Two Way MLR 5 day	1.3	HPAEC TNF alpha + IL-1 beta	100.0
Two Way MLR 7 day	0.0	Lung fibroblast none	2.9
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	3.4
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	1.9
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	5.4
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	11.3
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	49.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	6.7
Dendritic cells none	4.9	Dermal fibroblast IL-4	0.6
Dendritic cells LPS	4.3	Dermal Fibroblasts rest	3.5
Dendritic cells anti-CD40	8.7	Neutrophils TNFa+LPS	0.9
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	3.9	Colon	0.9
Macrophages rest	0.0	Lung	27.4
Macrophages LPS	4.7	Thymus	17.0
HUVEC none	0.0	Kidney	81.8
HUVEC starved	0.0		

Table RM. Panel 4D

Tissue Name	Rel. Exp.(%) Ag1860, Run 165828919	Rel. Exp.(%) Ag3112, Run 164526081	Tissue Name	Rel. Exp.(%) Ag1860, Run 165828919	Rel. Exp.(%) Ag3112, Run 164526081
Secondary Th1 act	0.0	0.0	HUVEC IL-1beta	6.5	2.2
Secondary Th2 act	0.0	0.0	HUVEC IFN gamma	0.0	0.0
Secondary Tr1 act	0.8	0.0	HUVEC TNF alpha + IFN gamma	0.6	0.0
Secondary Th1 rest	0.0	1.7	HUVEC TNF alpha + IL4	0.6	0.0
Secondary Th2 rest	0.0	0.0	HUVEC IL-11	0.0	0.0
Secondary Tr1 rest	0.6	0.0	Lung Microvascular EC none	0.0	0.0
Primary Th1 act	0.0	0.0	Lung Microvascular EC TNFalpha + IL- 1beta	9.7	9.7
Primary Th2 act	0.0	0.0	Microvascular Dermal EC none	0.0	0.0
Primary Tr1 act	0.7	0.0	Microvascular Dermal EC TNFalpha + IL- 1beta	1.4	5.9
Primary Th1 rest	0.7	0.0	Bronchial epithelium TNFalpha + IL1beta	0.6	0.0
Primary Th2 rest	0.6	0.0	Small airway epithelium none	0.0	0.0
Primary Tr1 rest	0.0	0.0	Small airway epithelium TNFalpha + IL- 1beta	0.0	0.0
CD45RA CD4 lymphocyte act	6.7	2.9	Coronary artery SMC rest	9.7	10.4
CD45RO CD4 lymphocyte act	0.0	0.0	Coronary artery SMC TNFalpha + IL-1beta	23.7	32.8
CD8 lymphocyte act	1.4	0.0	Astrocytes rest	0.0	0.0
Secondary CD8 lymphocyte rest	0.0	0.0	Astrocytes TNFalpha + IL- 1beta	11.7	0.0

Secondary CD8 lymphocyte act	0.0	1.9	KU-812 (Basophil) rest	0.0	0.0
CD4 lymphocyte none	0.0	0.0	KU-812 (Basophil) PMA/ionomycin	0.0	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	2.0	CCD1106 (Keratinocytes) none	0.0	0.0
LAK cells rest	0.9	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0	0.0
LAK cells IL-2	2.7	0.0	Liver cirrhosis	20.0	6.4
LAK cells IL-2+IL-12	0.0	0.0	Lupus kidney	0.0	0.0
LAK cells IL-2+IFN gamma	0.4	0.0	NCI-H292 none	0.0	0.0
LAK cells IL-2+IL-18	0.0	0.0	NCI-H292 IL-4	0.0	0.0
LAK cells PMA/ionomycin	1.8	0.0	NCI-H292 IL-9	0.0	2.6
NK Cells IL-2 rest	1.7	0.0	NCI-H292 IL-13	0.0	1.7
Two Way MLR 3 day	1.6	0.0	NCI-H292 IFN gamma	0.0	0.0
Two Way MLR 5 day	0.0	0.0	HPAEC none	0.0	0.0
Two Way MLR 7 day	0.0	0.0	HPAEC TNF alpha + IL-1 beta	100.0	69.3
PBMC rest	0.8	0.0	Lung fibroblast none	0.0	0.0
PBMC PWM	1.5	1.2	Lung fibroblast TNF alpha + IL-1 beta	16.4	3.5
PBMC PHA-L	0.0	0.0	Lung fibroblast IL-4	0.0	0.0
Ramos (B cell) none	0.0	0.0	Lung fibroblast IL-9	0.0	0.0
Ramos (B cell) ionomycin	0.0	0.0	Lung fibroblast IL-13	0.7	0.0
B lymphocytes PWM	0.0	2.5	Lung fibroblast IFN gamma	0.0	0.0
B lymphocytes CD40L and IL-4	0.0	1.7	Dermal fibroblast CCD1070 rest	2.3	4.1
EOL-1 dbcAMP	0.0	0.0	Dermal fibroblast CCD1070 TNF alpha	28.7	42.6

EOL-1 dbcAMP PMA/ionomycin	0.0	0.0	Dermal fibroblast CCD1070 IL-1 beta	49.7	100.0
Dendritic cells none	4.6	0.0	Dermal fibroblast. IFN gamma	0.0	0.9
Dendritic cells LPS	4.2	2.9	Dermal fibroblast IL-4	0.3	0.0
Dendritic cells anti-CD40	15.8	11.5	IBD Colitis 2	1.9	0.0
Monocytes rest	0.0	0.0	IBD Crohn's	2.7	5.1
Monocytes LPS	2.3	0.0	Colon	15.6	3.2
Macrophages rest	5.7	3.4	Lung	23.0	27.7
Macrophages LPS	6.8	3.4	Thymus	0.0	1.7
HUVEC none	0.0	0.0	Kidney	26.8	27.5
HUVEC starved	0.0	0.0			

Table RN. Panel CNS_1

Tissue Name	Rel. Exp.(%) Ag1860, Run 171634856	Tissue Name	Rel. Exp.(%) Ag1860, Run 171634856
BA4 Control	11.7	BA17 PSP	7.5
BA4 Control2	15.2	BA17 PSP2	7.9
BA4 Alzheimer's2	4.8	Sub Nigra Control	0.0
BA4 Parkinson's	23.3	Sub Nigra Control2	1.4
BA4 Parkinson's2	18.6	Sub Nigra Alzheimer's2	0.0
BA4 Huntington's	4.9	Sub Nigra Parkinson's2	2.3
BA4 Huntington's2	2.5	Sub Nigra Huntington's	10.5
BA4 PSP	5.0	Sub Nigra Huntington's2	23.3
BA4 PSP2	2.5	Sub Nigra PSP2	1.2
BA4 Depression	5.9	Sub Nigra Depression	0.0
BA4 Depression2	3.3	Sub Nigra Depression2	5.1
BA7 Control	19.5	Glob Palladus Control	12.1
BA7 Control2	23.5	Glob Palladus Control2	16.6
BA7 Alzheimer's2	10.5	Glob Palladus Alzheimer's	6.8
BA7 Parkinson's	11.0	Glob Palladus Alzheimer's2	11.0
BA7	15.5	Glob Palladus	100.0

Parkinson's2		Parkinson's	
BA7 Huntington's	11.8	Glob Palladus Parkinson's2	30.6
BA7 Huntington's2	9.5	Glob Palladus PSP	1.7
BA7 PSP	2.3	Glob Palladus PSP2	21.9
BA7 PSP2	5.5	Glob Palladus Depression	1.5
BA7 Depression	6.0	Temp Pole Control	11.1
BA9 Control	14.8	Temp Pole Control2	47.3
BA9 Control2	33.4	Temp Pole Alzheimer's	2.1
BA9 Alzheimer's	6.7	Temp Pole Alzheimer's2	4.6
BA9 Alzheimer's2	11.1	Temp Pole Parkinson's	19.1
BA9 Parkinson's	11.8	Temp Pole Parkinson's2	20.9
BA9 Parkinson's2	24.1	Temp Pole Huntington's	25.5
BA9 Huntington's	18.8	Temp Pole PSP	0.0
BA9 Huntington's2	8.0	Temp Pole PSP2	7.7
BA9 PSP	0.0	Temp Pole Depression2	1.9
BA9 PSP2	3.8	Cing Gyr Control	4.7
BA9 Depression	3.5	Cing Gyr Control2	20.4
BA9 Depression2	5.6	Cing Gyr Alzheimer's	6.0
BA17 Control	21.3	Cing Gyr Alzheimer's2	4.5
BA17 Control2	46.0	Cing Gyr Parkinson's	1.7
BA17 Alzheimer's2	4.0	Cing Gyr Parkinson's2	10.4
BA17 Parkinson's	12.5	Cing Gyr Huntington's	20.2
BA17 Parkinson's2	35.6	Cing Gyr Huntington's2	1.9
BA17 Huntington's	18.6	Cing Gyr PSP	0.0
BA17 Huntington's2	10.2	Cing Gyr PSP2	2.3
BA17 Depression	11.2	Cing Gyr Depression	0.0

BA17 Depression2	18.4	Cing Gyr Depression2	1.4
---------------------	------	----------------------	-----

AI_comprehensive_panel_v1.0 Summary: Ag1860 Highest expression of the CG57109-01 transcript in this panel is seen in synovium from an OA patient (CT=33.7). Overall, this gene is expressed in OA tissue but not in normal joint tissue and is expressed in pulmonary tissue from patients with atopic asthma but not in normal lung tissue. Please
5 see panel 4D for discussion of utility of this gene in inflammation.

CNS_neurodegeneration_v1.0 Summary: Ag1860/Ag3112/Ag4281 Three experiments with two different probe and primer sets produce results that are in very good agreement. This panel does not show differential expression of the CG57109-01 gene in Alzheimer's disease. However, this expression profile confirms the presence of this gene
10 in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system. One experiment with Ag6051, which is specific to CG57109-06 only, is in general agreement with the results above. Highest expression in this panel was in the temporal cortex of an Alzheimer's patient (CT=34.6).

General_screening_panel_v1.4 Summary: Ag4281 Highest expression of the
15 CG57109-01 gene appears to be in the fetal brain (CT=29.5). Overall, expression of this gene appears to be highly brain-specific in this panel. with moderate levels of expression in the amygdala, hippocampus, thalamus and spinal cord and low but significant levels in the cerebral cortex and the substantia nigra. This gene encodes a novel doublecortin/CAM
20 kinase like protein. Other members of this family have been implicated in the calcium-signaling pathway that controls neuronal migration in the developing brain. In addition, CAM kinase has been shown to play a crucial role in hippocampal LTP from studies in transgenic and knock-out mice, and may also play a role in memory formation in the mature nervous system as well as the developing brain. CAM kinases have also been
25 shown to phosphorylate tau, an integral component of the neurofibrillary tangles seen in Alzheimer's, in a manner which shifts tau electrophoretic motility to that seen in the AD brain. Furthermore, tau from AD brains shows aberrant phosphorylation. Thus, based on the expression of this DCAM kinase homolog in the brain, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of learning and
30 memory deficits that are a result of aging or neurodegenerative disease and also in the treatment of neurologic disorders themselves, including Alzheimer's disease.

Moderate to low levels of expression are also seen in a variety of samples from normal tissues, including testis, fetal and adult heart and skeletal muscle and fetal lung.

In addition, this gene is expressed at much higher levels in fetal lung (CT=32.3) when compared to expression in the adult counterpart (CT=40). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue.

Panel 1.3D Summary: Ag1860/Ag3112 Two experiments with the same probe and primer set produce results that are in excellent agreement, with highest expression of the CG57109-01 gene in the spinal cord (CTs=31.7). Expression of this gene is restricted to the nervous system and the testis. Thus, expression of this gene could be used to differentiate between neural and non-neural tissue. Please see Panel 1.4 for further discussion of utility of this gene in the CNS.

Panel 2.2 Summary: Ag1860 Expression of the CG57109-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

Panel 3D Summary: Ag3112 Expression of the CG57109-01 gene is restricted to a sample derived from a lung cancer cell line (CT=32.6). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of lung cancer. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of lung cancer.

Panel 4.1D Summary: Ag1860 This CG57109-01 transcript is highly expressed in activated dermal fibroblasts, endothelial cells, and astrocytes after treatment with IL-1 or TNFalpha. Highest expression is seen in treated HPAECs (CT=31.3). Please see panel 4D for discussion of utility of this gene in inflammation.

Panel 4D Summary: Ag1860 This transcript is highly expressed in activated dermal fibroblasts, endothelial cells, and astrocytes after treatment with IL-1 or TNFalpha, with highest expression in TNF alpha and IL-1 beta treated HPAECs (CT=30.9). This protein has homology to protein kinase and may be involved in leukocyte extravasation from the peripheral blood into tissues. (Borbiev T, Am J Physiol Lung Cell Mol Physiol 2001 May;280(5):L983-90) Therefore, antagonistic therapeutics designed against the protein encoded by this transcript may reduce or inhibit inflammation due to asthma, allergy, emphysema, osteoarthritis, colitis, psoriasis, or delayed type hypersensitivity. Agonistic therapies may also direct leukocyte traffic into tumors or sites of infection.

Ag3112 Highest expression of the transcript is seen in IL-1 beta treated dermal fibroblasts (CT=30.4). Expression is in agreement with the profile seen with Ag1860, except no expression is seen in astrocytes.

Panel CNS_1 Summary: Ag1860 This panel confirms expression of the CG57109-01 gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system.

General oncology screening panel_v_2.4 Summary: Ag1860 Expression of the CG57109-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.) Ag6051 Expression of the CG57109-06 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

S. CG57366-01: Kiaa1223

Expression of gene CG57366-01 was assessed using the primer-probe set Ag3219, described in Table SA. Results of the RTQ-PCR runs are shown in Tables SB, SC, SD and SE.

Table SA. Probe Name Ag3219

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agggatacccttcctcagaga-3'	21	3380	272
Probe	TET-5'-agccttaaacaagctctgaagcttca-3'-TAMRA	26	3409	273
Reverse	5'-gctagggtcagaaccttcaatc-3'	22	3435	274

Table SB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3219, Run 209862297	Rel. Exp.(%) Ag3219, Run 224079667	Tissue Name	Rel. Exp.(%) Ag3219, Run 209862297	Rel. Exp.(%) Ag3219, Run 224079667
AD 1 Hippo	18.9	14.7	Control (Path) 3 Temporal Ctx	9.5	9.2
AD 2 Hippo	41.2	52.5	Control (Path) 4 Temporal Ctx	47.3	41.2
AD 3 Hippo	1.9	14.0	AD 1 Occipital Ctx	20.4	28.5
AD 4 Hippo	6.3	9.9	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 Hippo	85.3	90.1	AD 3 Occipital Ctx	5.0	10.7
AD 6 Hippo	54.3	64.2	AD 4 Occipital Ctx	28.9	19.5

Control 2 Hippo	30.1	39.8	AD 5 Occipital Ctx	32.8	21.5
Control 4 Hippo	15.7	18.2	AD 6 Occipital Ctx	24.1	40.9
Control (Path) 3 Hippo	11.8	12.6	Control 1 Occipital Ctx	5.8	6.7
AD 1 Temporal Ctx	0.4	36.3	Control 2 Occipital Ctx	36.9	44.4
AD 2 Temporal Ctx	31.6	42.3	Control 3 Occipital Ctx	22.8	21.3
AD 3 Temporal Ctx	2.3	8.6	Control 4 Occipital Ctx	8.2	11.8
AD 4 Temporal Ctx	19.9	30.1	Control (Path) 1 Occipital Ctx	88.9	79.6
AD 5 Inf Temporal Ctx	99.3	100.0	Control (Path) 2 Occipital Ctx	19.9	21.0
AD 5 Sup Temporal Ctx	55.9	71.7	Control (Path) 3 Occipital Ctx	3.6	7.0
AD 6 Inf Temporal Ctx	52.5	66.0	Control (Path) 4 Occipital Ctx	21.6	26.6
AD 6 Sup Temporal Ctx	58.2	65.5	Control 1 Parietal Ctx	13.2	13.5
Control 1 Temporal Ctx	9.5	10.7	Control 2 Parietal Ctx	0.0	65.1
Control 2 Temporal Ctx	32.8	37.1	Control 3 Parietal Ctx	21.9	18.0
Control 3 Temporal Ctx	23.7	24.8	Control (Path) 1 Parietal Ctx	100.0	88.3
Control 3 Temporal Ctx	9.1	11.2	Control (Path) 2 Parietal Ctx	27.2	24.3
Control (Path) 1 Temporal	83.5	63.3	Control (Path) 3 Parietal Ctx	5.4	7.5

Ctx					
Control (Path) 2 Temporal Ctx	50.3	52.5	Control (Path) 4 Parietal Ctx	53.6	48.0

Table SC. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag3219. Run 168013879	Tissue Name	Rel. Exp.(%) Ag3219, Run 168013879
Liver adenocarcinoma	14.9	Kidney (fetal)	31.0
Pancreas	1.8	Renal ca. 786-0	19.3
Pancreatic ca. CAPAN-2	17.4	Renal ca. A498	10.4
Adrenal gland	1.7	Renal ca. RXF 393	8.5
Thyroid	5.3	Renal ca. ACHN	12.3
Salivary gland	2.9	Renal ca. UO-31	14.4
Pituitary gland	2.4	Renal ca. TK-10	16.8
Brain (fetal)	58.2	Liver	3.0
Brain (whole)	12.9	Liver (fetal)	3.7
Brain (amygdala)	10.6	Liver ca. (hepatoblast) HepG2	6.2
Brain (cerebellum)	9.7	Lung	4.5
Brain (hippocampus)	8.4	Lung (fetal)	27.7
Brain (substantia nigra)	8.5	Lung ca. (small cell) LX-1	5.3
Brain (thalamus)	5.3	Lung ca. (small cell) NCI-H69	17.1
Cerebral Cortex	7.1	Lung ca. (s.cell var.) SHP-77	42.9
Spinal cord	5.5	Lung ca. (large cell)NCI-H460	4.0
glio/astro U87-MG	9.9	Lung ca. (non-sm. cell) A549	22.7
glio/astro U-118-MG	18.2	Lung ca. (non-s.cell) NCI-H23	17.0
astrocytoma SW1783	17.0	Lung ca. (non-s.cell) HOP-62	17.2
neuro*; met SK-N-AS	28.1	Lung ca. (non-s.cl) NCI-H522	14.4
astrocytoma SF-539	18.4	Lung ca. (squam.) SW 900	65.1
astrocytoma SNB-75	44.1	Lung ca. (squam.) NCI-	57.4

		H596	
glioma SNB-19	14.5	Mammary gland	7.5
glioma U251	55.9	Breast ca.* (pl.ef) MCF-7	46.3
glioma SF-295	21.2	Breast ca.* (pl.ef) MDA-MB-231	16.3
Heart (fetal)	2.2	Breast ca.* (pl.ef) T47D	31.9
Heart	3.9	Breast ca. BT-549	9.5
Skeletal muscle (fetal)	1.6	Breast ca. MDA-N	10.4
Skeletal muscle	3.1	Ovary	3.8
Bone marrow	1.5	Ovarian ca. OVCAR-3	17.6
Thymus	6.4	Ovarian ca. OVCAR-4	5.4
Spleen	1.1	Ovarian ca. OVCAR-5	80.1
Lymph node	2.8	Ovarian ca. OVCAR-8	9.3
Colorectal	4.0	Ovarian ca. IGROV-1	4.9
Stomach	5.0	Ovarian ca.* (ascites) SK-OV-3	100.0
Small intestine	3.7	Uterus	7.2
Colon ca. SW480	2.8	Placenta	1.4
Colon ca.* SW620(SW480 met)	9.5	Prostate	4.8
Colon ca. HT29	9.5	Prostate ca.* (bone met)PC-3	29.3
Colon ca. HCT-116	8.3	Testis	8.2
Colon ca. CaCo-2	5.6	Melanoma Hs688(A).T	8.9
Colon ca. tissue(ODO3866)	5.4	Melanoma* (met) Hs688(B).T	13.4
Colon ca. HCC-2998	8.7	Melanoma UACC-62	3.0
Gastric ca.* (liver met) NCI-N87	16.5	Melanoma M14	2.8
Bladder	21.3	Melanoma LOX IMVI	11.0
Trachea	2.3	Melanoma* (met) SK- MEL-5	7.5
Kidney	6.3	Adipose	17.3

Table SD. Panel 2.2

Tissue Name	Rel. Exp.(%) Ag3219, Run 174416266	Tissue Name	Rel. Exp.(%) Ag3219, Run 174416266
Normal Colon	22.8	Kidney Margin (OD04348)	51.4
Colon cancer	25.2	Kidney malignant cancer	36.6

(OD06064)		(OD06204B)	
Colon Margin (OD06064)	12.0	Kidney normal adjacent tissue (OD06204E)	6.3
Colon cancer (OD06159)	2.3	Kidney Cancer (OD04450-01)	54.3
Colon Margin (OD06159)	14.5	Kidney Margin (OD04450-03)	9.7
Colon cancer (OD06297-04)	4.1	Kidney Cancer 8120613	0.0
Colon Margin (OD06297-05)	23.0	Kidney Margin 8120614	0.9
CC Gr.2 ascend colon (ODO3921)	3.6	Kidney Cancer 9010320	3.0
CC Margin (ODO3921)	6.9	Kidney Margin 9010321	3.0
Colon cancer metastasis (OD06104)	3.5	Kidney Cancer 8120607	2.6
Lung Margin (OD06104)	3.6	Kidney Margin 8120608	0.1
Colon mets to lung (OD04451-01)	14.3	Normal Uterus	44.1
Lung Margin (OD04451-02)	29.5	Uterine Cancer 064011	13.2
Normal Prostate	5.8	Normal Thyroid	3.0
Prostate Cancer (OD04410)	11.3	Thyroid Cancer 064010	10.2
Prostate Margin (OD04410)	12.2	Thyroid Cancer A302152	14.0
Normal Ovary	2.4	Thyroid Margin A302153	5.1
Ovarian cancer (OD06283-03)	2.7	Normal Breast	25.2
Ovarian Margin (OD06283-07)	14.0	Breast Cancer (OD04566)	19.6
Ovarian Cancer 064008	9.9	Breast Cancer 1024	8.5
Ovarian cancer (OD06145)	7.1	Breast Cancer (OD04590-01)	17.1
Ovarian Margin (OD06145)	23.0	Breast Cancer Mets (OD04590-03)	31.2
Ovarian cancer (OD06455-03)	17.3	Breast Cancer Metastasis (OD04655-05)	100.0
Ovarian Margin (OD06455-07)	12.2	Breast Cancer 064006	27.9

Normal Lung	21.2	Breast Cancer 9100266	9.8
Invasive poor diff. lung adeno (ODO4945-01)	20.7	Breast Margin 9100265	13.2
Lung Margin (ODO4945-03)	24.1	Breast Cancer A209073	14.1
Lung Malignant Cancer (OD03126)	8.7	Breast Margin A2090734	22.4
Lung Margin (OD03126)	16.2	Breast cancer (OD06083)	100.0
Lung Cancer (OD05014A)	17.0	Breast cancer node metastasis (OD06083)	60.7
Lung Margin (OD05014B)	42.3	Normal Liver	12.8
Lung cancer (OD06081)	10.4	Liver Cancer 1026	0.4
Lung Margin (OD06081)	13.2	Liver Cancer 1025	3.8
Lung Cancer (OD04237-01)	10.4	Liver Cancer 6004-T	1.1
Lung Margin (OD04237-02)	30.6	Liver Tissue 6004-N	0.5
Ocular Melanoma Metastasis	9.8	Liver Cancer 6005-T	2.6
Ocular Melanoma Margin (Liver)	7.3	Liver Tissue 6005-N	4.2
Melanoma Metastasis	32.5	Liver Cancer 064003	7.1
Melanoma Margin (Lung)	21.6	Normal Bladder	18.6
Normal Kidney	4.7	Bladder Cancer 1023	3.5
Kidney Ca, Nuclear grade 2 (OD04338)	28.3	Bladder Cancer A302173	10.8
Kidney Margin (OD04338)	11.8	Normal Stomach	27.4
Kidney Ca Nuclear grade 1/2 (OD04339)	34.2	Gastric Cancer 9060397	2.4
Kidney Margin (OD04339)	8.2	Stomach Margin 9060396	3.1
Kidney Ca, Clear cell type (OD04340)	11.0	Gastric Cancer 9060395	5.2
Kidney Margin (OD04340)	13.6	Stomach Margin 9060394	4.3
Kidney Ca, Nuclear grade 3 (OD04348)	4.5	Gastric Cancer 064005	5.8

Table SE. Panel 4D

Tissue Name	Rel. Exp.(%) Ag3219, Run 164531987	Tissue Name	Rel. Exp.(%) Ag3219, Run 164531987
Secondary Th1 act	0.1	HUVEC IL-1beta	27.5
Secondary Th2 act	0.1	HUVEC IFN gamma	35.8
Secondary Tr1 act	0.3	HUVEC TNF alpha + IFN gamma	39.5
Secondary Th1 rest	0.3	HUVEC TNF alpha + IL4	54.7
Secondary Th2 rest	0.4	HUVEC IL-11	30.4
Secondary Tr1 rest	0.3	Lung Microvascular EC none	63.7
Primary Th1 act	0.5	Lung Microvascular EC TNFalpha + IL-1beta	43.2
Primary Th2 act	0.6	Microvascular Dermal EC none	49.3
Primary Tr1 act	0.6	Microvascular Dermal EC TNFalpha + IL-1beta	19.5
Primary Th1 rest	1.6	Bronchial epithelium TNFalpha + IL1beta	24.3
Primary Th2 rest	1.2	Small airway epithelium none	6.7
Primary Tr1 rest	0.5	Small airway epithelium TNFalpha + IL-1beta	50.0
CD45RA CD4 lymphocyte act	6.6	Coronary artery SMC rest	39.0
CD45RO CD4 lymphocyte act	0.2	Coronary artery SMC TNFalpha + IL-1beta	18.6
CD8 lymphocyte act	0.1	Astrocytes rest	10.2
Secondary CD8 lymphocyte rest	0.4	Astrocytes TNFalpha + IL-1beta	7.9
Secondary CD8 lymphocyte act	0.1	KU-812 (Basophil) rest	13.3
CD4 lymphocyte none	0.1	KU-812 (Basophil) PMA/ionomycin	42.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.2	CCD1106 (Keratinocytes) none	16.2
LAK cells rest	2.3	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	3.8
LAK cells IL-2	0.4	Liver cirrhosis	2.1
LAK cells IL-2+IL-12	0.2	Lupus kidney	1.3
LAK cells IL-2+IFN	0.9	NCI-H292 none	16.6

gamma			
LAK cells IL-2+ IL-18	0.5	NCI-H292 IL-4	22.8
LAK cells PMA/ionomycin	1.2	NCI-H292 IL-9	19.8
NK Cells IL-2 rest	0.1	NCI-H292 IL-13	14.5
Two Way MLR 3 day	1.0	NCI-H292 IFN gamma	14.2
Two Way MLR 5 day	0.7	HPAEC none	43.8
Two Way MLR 7 day	0.3	HPAEC TNF alpha + IL-1 beta	24.0
PBMC rest	1.8	Lung fibroblast none	8.9
PBMC PWM	2.2	Lung fibroblast TNF alpha + IL-1 beta	9.5
PBMC PHA-L	0.8	Lung fibroblast IL-4	16.5
Ramos (B cell) none	2.9	Lung fibroblast IL-9	12.6
Ramos (B cell) ionomycin	15.3	Lung fibroblast IL-13	11.9
B lymphocytes PWM	4.4	Lung fibroblast IFN gamma	10.4
B lymphocytes CD40L and IL-4	0.7	Dermal fibroblast CCD1070 rest	28.1
EOL-1 dbcAMP	3.7	Dermal fibroblast CCD1070 TNF alpha	22.2
EOL-1 dbcAMP PMA/ionomycin	2.0	Dermal fibroblast CCD1070 IL-1 beta	16.4
Dendritic cells none	4.6	Dermal fibroblast IFN gamma	8.7
Dendritic cells LPS	3.2	Dermal fibroblast IL-4	29.5
Dendritic cells anti- CD40	4.4	IBD Colitis 2	1.3
Monocytes rest	12.8	IBD Crohn's	1.7
Monocytes LPS	2.4	Colon	9.7
Macrophages rest	9.2	Lung	27.2
Macrophages LPS	1.1	Thymus	14.2
HUVEC none	57.4	Kidney	22.8
HUVEC starved	100.0		

5 **CNS_neurodegeneration_v1.0 Summary:** Ag3219 Two experiments with the same probe and primer set produce results that are in excellent agreement. This gene is downregulated in the temporal cortex of Alzheimer's disease patients when compared with non-demented controls ($p = 0.001$ when analyzed by Ancova, estimate of total cDNA loaded per well used as a covariate). Therefore, up-regulation of this gene or its protein

product, or treatment with specific agonists for this receptor may be of use in reversing the dementia/memory loss and neuronal death associated with this disease.

Panel 1.3D Summary: Ag3219 Highest expression of the CG57366-01 gene is seen in an ovarian cancer cell line (CT=29.9). In addition, expression appears to be
5 significant in all the cell lines on this panel. In addition, this gene appears to be expressed in most of the samples on this panel, suggesting a role for this gene product in cell survival and proliferation. Thus, expression of this gene could be used as a marker of ovarian cancer. Furthermore, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of cancer.

10 Among tissues with metabolic function, this gene is expressed at moderate to low levels in pituitary, adipose, thyroid, skeletal muscle, heart, and adult and fetal liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity
15 and diabetes.

This gene is also expressed at moderate to low levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease,
20 schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 2.2 Summary: Ag3219 Highest expression of the CG57366-01 gene is seen in breast cancer (CT=29.3). In addition, significant expression is seen in a cluster of breast cancer samples. Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of breast
25 cancer. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of breast cancer.

Panel 4D Summary: Ag3219 Highest expression of the CG57366-01 gene is seen in untreated HUVEC (CT=27.7). Expression on this panel is seen mainly in endothelial cells and fibroblasts from lung and skin, basophils and astrocytes, and normal lung,
30 thymus and kidney. Thus, this gene product may be involved in pathological and inflammatory lung and skin conditions, including asthma, emphysema, allergy and psoriasis.

T. CG57368-01: Adenylate Cyclase Type IV

Expression of gene CG57368-01 was assessed using the primer-probe set Ag3221, described in Table TA. Results of the RTQ-PCR runs are shown in Tables TB, TC, TD and TE.

5

Table TA. Probe Name Ag3221

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - aactgatgggtgctatctcctt - 3'	22	2390	275
Probe	TET-5' - cttctttcttcacctccttgctcctgg - 3' - TAMRA	26	2418	276
Reverse	5' - aggcggcagtagtactcattct - 3'	22	2450	277

Table TB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3221, Run 209862298	Tissue Name	Rel. Exp.(%) Ag3221, Run 209862298
AD 1 Hippo	4.2	Control (Path) 3 Temporal Ctx	17.2
AD 2 Hippo	41.5	Control (Path) 4 Temporal Ctx	42.9
AD 3 Hippo	8.0	AD 1 Occipital Ctx	1.7
AD 4 Hippo	13.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	56.6	AD 3 Occipital Ctx	5.3
AD 6 Hippo	49.0	AD 4 Occipital Ctx	18.9
Control 2 Hippo	13.5	AD 5 Occipital Ctx	33.0
Control 4 Hippo	21.9	AD 6 Occipital Ctx	32.1
Control (Path) 3 Hippo	14.7	Control 1 Occipital Ctx	57.8
AD 1 Temporal Ctx	12.6	Control 2 Occipital Ctx	40.3
AD 2 Temporal Ctx	20.3	Control 3 Occipital Ctx	23.2
AD 3 Temporal Ctx	5.1	Control 4 Occipital Ctx	10.2
AD 4 Temporal Ctx	32.8	Control (Path) 1 Occipital Ctx	87.7
AD 5 Inf Temporal Ctx	62.9	Control (Path) 2 Occipital Ctx	40.1
AD 5 SupTemporal Ctx	47.3	Control (Path) 3 Occipital Ctx	25.5
AD 6 Inf	56.3	Control (Path) 4	71.7

Temporal Ctx		Occipital Ctx	
AD 6 Sup Temporal Ctx	65.5	Control 1 Parietal Ctx	24.0
Control 1 Temporal Ctx	33.9	Control 2 Parietal Ctx	67.8
Control 2 Temporal Ctx	27.2	Control 3 Parietal Ctx	18.0
Control 3 Temporal Ctx	11.9	Control (Path) 1 Parietal Ctx	43.8
Control 4 Temporal Ctx	17.8	Control (Path) 2 Parietal Ctx	25.9
Control (Path) 1 Temporal Ctx	43.5	Control (Path) 3 Parietal Ctx	25.0
Control (Path) 2 Temporal Ctx	62.9	Control (Path) 4 Parietal Ctx	100.0

Table TC. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag3221, Run 168014001	Tissue Name	Rel. Exp.(%) Ag3221, Run 168014001
Liver adenocarcinoma	0.4	Kidney (fetal)	73.2
Pancreas	9.3	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.4	Renal ca. A498	0.7
Adrenal gland	14.9	Renal ca. RXF 393	0.0
Thyroid	12.2	Renal ca. ACHN	0.0
Salivary gland	6.5	Renal ca. UO-31	0.0
Pituitary gland	10.8	Renal ca. TK-10	0.0
Brain (fetal)	5.4	Liver	15.0
Brain (whole)	11.2	Liver (fetal)	3.0
Brain (amygdala)	10.2	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	12.1	Lung	54.7
Brain (hippocampus)	7.3	Lung (fetal)	94.0
Brain (substantia nigra)	11.7	Lung ca. (small cell) LX-1	2.1
Brain (thalamus)	11.3	Lung ca. (small cell) NCI-H69	1.5
Cerebral Cortex	10.3	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	17.1	Lung ca. (large cell)NCI-H460	0.0
glio/astro U87-MG	0.0	Lung ca. (non-sm. cell)	0.0

		A549	
glio/astro U-118-MG	0.0	Lung ca. (non-s.cell) NCI-H23	1.3
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP-62	0.0
neuro*; met SK-N-AS	0.0	Lung ca. (non-s.cl) NCI- H522	0.0
astrocytoma SF-539	0.0	Lung ca. (squam.) SW 900	1.5
astrocytoma SNB-75	0.7	Lung ca. (squam.) NCI- H596	0.0
glioma SNB-19	0.0	Mammary gland	100.0
glioma U251	0.0	Breast ca.* (pl.ef) MCF- 7	0.0
glioma SF-295	0.0	Breast ca.* (pl.ef) MDA- MB-231	0.0
Heart (fetal)	95.3	Breast ca.* (pl.ef) T47D	2.0
Heart	50.3	Breast ca. BT-549	0.0
Skeletal muscle (fetal)	57.8	Breast ca. MDA-N	0.0
Skeletal muscle	28.5	Ovary	18.2
Bone marrow	11.6	Ovarian ca. OVCAR-3	0.0
Thymus	19.3	Ovarian ca. OVCAR-4	0.0
Spleen	27.2	Ovarian ca. OVCAR-5	43.5
Lymph node	41.8	Ovarian ca. OVCAR-8	0.6
Colorectal	5.8	Ovarian ca. IGROV-1	0.0
Stomach	10.0	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	21.2	Uterus	62.9
Colon ca. SW480	0.0	Placenta	22.2
Colon ca.* SW620(SW480 met)	0.0	Prostate	10.4
Colon ca. HT29	0.4	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	5.0
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	6.7	Melanoma* (met) Hs688(B).T	5.5
Colon ca. HCC-2998	7.2	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	2.8	Melanoma M14	0.0
Bladder	8.5	Melanoma LOX IMVI	0.0

Trachea	6.1	Melanoma* (met) SK-MEL-5	0.0
Kidney	27.5	Adipose	42.3

Table TD. Panel 2.2

Tissue Name	Rel. Exp.(%) Ag3221, Run 173817111	Tissue Name	Rel. Exp.(%) Ag3221, Run 173817111
Normal Colon	8.0	Kidney Margin (OD04348)	100.0
Colon cancer (OD06064)	17.4	Kidney malignant cancer (OD06204B)	0.0
Colon Margin (OD06064)	10.2	Kidney normal adjacent tissue (OD06204E)	8.5
Colon cancer (OD06159)	2.5	Kidney Cancer (OD04450-01)	0.0
Colon Margin (OD06159)	4.2	Kidney Margin (OD04450-03)	18.4
Colon cancer (OD06297-04)	4.2	Kidney Cancer 8120613	5.2
Colon Margin (OD06297-05)	18.2	Kidney Margin 8120614	28.9
CC Gr.2 ascend colon (ODO3921)	6.1	Kidney Cancer 9010320	5.3
CC Margin (ODO3921)	1.6	Kidney Margin 9010321	4.4
Colon cancer metastasis (OD06104)	4.3	Kidney Cancer 8120607	8.0
Lung Margin (OD06104)	1.8	Kidney Margin 8120608	17.1
Colon mets to lung (OD04451-01)	8.6	Normal Uterus	47.3
Lung Margin (OD04451-02)	16.7	Uterine Cancer 064011	23.8
Normal Prostate	12.3	Normal Thyroid	6.0
Prostate Cancer (OD04410)	3.1	Thyroid Cancer 064010	2.7
Prostate Margin (OD04410)	7.0	Thyroid Cancer A302152	15.9
Normal Ovary	28.3	Thyroid Margin A302153	4.5
Ovarian cancer (OD06283-03)	6.0	Normal Breast	26.6
Ovarian Margin	44.1	Breast Cancer	3.8

(OD06283-07)		(OD04566)	
Ovarian Cancer 064008	4.3	Breast Cancer 1024	20.3
Ovarian cancer (OD06145)	2.3	Breast Cancer (OD04590-01)	4.7
Ovarian Margin (OD06145)	16.4	Breast Cancer Mets (OD04590-03)	14.5
Ovarian cancer (OD06455-03)	0.9	Breast Cancer Metastasis (OD04655-05)	11.0
Ovarian Margin (OD06455-07)	13.0	Breast Cancer 064006	18.8
Normal Lung	26.1	Breast Cancer 9100266	5.3
Invasive poor diff. lung adeno (ODO4945-01)	3.2	Breast Margin 9100265	4.4
Lung Margin (ODO4945-03)	36.6	Breast Cancer A209073	6.3
Lung Malignant Cancer (OD03126)	8.2	Breast Margin A2090734	18.2
Lung Margin (OD03126)	23.0	Breast cancer (OD06083)	12.2
Lung Cancer (OD05014A)	5.5	Breast cancer node metastasis (OD06083)	9.1
Lung Margin (OD05014B)	29.1	Normal Liver	6.1
Lung cancer (OD06081)	4.0	Liver Cancer 1026	4.7
Lung Margin (OD06081)	31.0	Liver Cancer 1025	13.0
Lung Cancer (OD04237-01)	1.5	Liver Cancer 6004-T	5.0
Lung Margin (OD04237-02)	35.1	Liver Tissue 6004-N	1.6
Ocular Melanoma Metastasis	0.0	Liver Cancer 6005-T	13.7
Ocular Melanoma Margin (Liver)	4.2	Liver Tissue 6005-N	7.8
Melanoma Metastasis	2.7	Liver Cancer 064003	3.5
Melanoma Margin (Lung)	26.8	Normal Bladder	12.9
Normal Kidney	11.5	Bladder Cancer 1023	6.5
Kidney Ca, Nuclear grade 2 (OD04338)	55.1	Bladder Cancer A302173	2.0

Kidney Margin (OD04338)	7.6	Normal Stomach	32.1
Kidney Ca Nuclear grade 1/2 (OD04339)	5.6	Gastric Cancer 9060397	1.0
Kidney Margin (OD04339)	14.8	Stomach Margin 9060396	8.5
Kidney Ca, Clear cell type (OD04340)	18.4	Gastric Cancer 9060395	10.9
Kidney Margin (OD04340)	9.3	Stomach Margin 9060394	16.0
Kidney Ca, Nuclear grade 3 (OD04348)	1.5	Gastric Cancer 064005	11.3

Table TE. Panel 4D

Tissue Name	Rel. Exp.(%) Ag3221, Run 164531988	Tissue Name	Rel. Exp.(%) Ag3221, Run 164531988
Secondary Th1 act	0.1	HUVEC IL-1beta	2.6
Secondary Th2 act	0.2	HUVEC IFN gamma	25.5
Secondary Tr1 act	0.4	HUVEC TNF alpha + IFN gamma	39.2
Secondary Th1 rest	1.0	HUVEC TNF alpha + IL4	31.9
Secondary Th2 rest	0.7	HUVEC IL-11	21.3
Secondary Tr1 rest	1.3	Lung Microvascular EC none	48.0
Primary Th1 act	0.2	Lung Microvascular EC TNFalpha + IL-1beta	100.0
Primary Th2 act	0.0	Microvascular Dermal EC none	47.3
Primary Tr1 act	0.6	Microvascular Dermal EC TNFalpha + IL-1beta	44.1
Primary Th1 rest	1.6	Bronchial epithelium TNFalpha + IL1beta	6.8
Primary Th2 rest	2.6	Small airway epithelium none	1.7
Primary Tr1 rest	1.7	Small airway epithelium TNFalpha + IL-1beta	6.9
CD45RA CD4 lymphocyte act	3.3	Coronary artery SMC rest	1.7
CD45RO CD4 lymphocyte act	0.2	Coronary artery SMC TNFalpha + IL-1beta	2.1
CD8 lymphocyte act	0.4	Astrocytes rest	0.0
Secondary CD8	0.5	Astrocytes TNFalpha +	0.0

lymphocyte rest		IL-1beta	
Secondary CD8 lymphocyte act	0.2	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	1.1	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	3.5	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.7
LAK cells IL-2	2.4	Liver cirrhosis	0.6
LAK cells IL-2+IL-12	0.0	Lupus kidney	6.5
LAK cells IL-2+IFN gamma	0.6	NCI-H292 none	0.5
LAK cells IL-2+ IL-18	0.5	NCI-H292 IL-4	0.0
LAK cells PMA/ionomycin	0.2	NCI-H292 IL-9	0.5
NK Cells IL-2 rest	0.8	NCI-H292 IL-13	0.0
Two Way MLR 3 day	0.7	NCI-H292 IFN gamma	0.0
Two Way MLR 5 day	0.0	HPAEC none	45.1
Two Way MLR 7 day	0.2	HPAEC TNF alpha + IL-1 beta	41.2
PBMC rest	1.1	Lung fibroblast none	1.5
PBMC PWM	0.5	Lung fibroblast TNF alpha + IL-1 beta	0.2
PBMC PHA-L	0.5	Lung fibroblast IL-4	0.2
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	0.3
B lymphocytes PWM	0.4	Lung fibroblast IFN gamma	0.2
B lymphocytes CD40L and IL-4	0.9	Dermal fibroblast CCD1070 rest	4.0
EOL-1 dbcAMP	1.3	Dermal fibroblast CCD1070 TNF alpha	8.0
EOL-1 dbcAMP PMA/ionomycin	2.6	Dermal fibroblast CCD1070 IL-1 beta	5.3
Dendritic cells none	1.3	Dermal fibroblast IFN gamma	7.1
Dendritic cells LPS	1.0	Dermal fibroblast IL-4	9.7
Dendritic cells anti-CD40	3.7	IBD Colitis 2	0.7

Monocytes rest	5.7	IBD Crohn's	3.0
Monocytes LPS	4.5	Colon	9.4
Macrophages rest	12.4	Lung	17.0
Macrophages LPS	1.4	Thymus	14.0
HUVEC none	17.9	Kidney	7.4
HUVEC starved	16.6		

CNS_neurodegeneration_v1.0 Summary: Ag3221 This panel confirms the expression of this gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please
5 see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

Panel 1.3D Summary: Ag3221 Highest expression of the CG57368-01 gene is detected in mammary gland (CT=30). High expression of this gene is seen mainly in the normal tissue samples suggesting an important role in cellular function.

10 Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

15 In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

20 **Panel 2.2 Summary:** Ag3221 Highest expression of the CG57368-01 gene is detected in control kidney (OD04348) sample (CT=32.7). Expression of this gene is generally higher in the normal control margin samples as compared to the cancer tissues. Interestingly, expression of this gene is higher in kidney cancer nuclear grade 2 (OD04338) (Ct=33.5) as compared to corresponding control sample (Ct=36.4). Therefore,
25 expression of this gene can be used as a diagnostic marker for nuclear grade 2 kidney cancer and therapeutic modulation of this gene could be beneficial in the treatment of kidney cancer.

Panel 4D Summary: Ag3221 Highest expression of the CG57368-01 gene is detected in TNFalpha + IL-1beta treated lung microvascular EC (CT=28). In addition, moderate to low expression of the gene is also seen in resting primary and secondary Th1,Th2,Tr1, B lymphocytes, LAK cells, dendritic cells, monocytes, macrophages, endothelial cells, eosinophils, small airway epithelium, dermal fibroblasts, lupus kidney and normal tissues represented by colon, lung, thymus and kidney. Therefore, therapeutic modulation of this gene through the use of small molecule drug may be useful in the treatment of autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis

Panel CNS_1 Summary: Ag3221 Results from one experiment with the CG57368-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

15 U. CG59955-01: GPCR

Expression of gene CG59955-01 was assessed using the primer-probe set Ag3637, described in Table UA.

Table UA. Probe Name Ag3637

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-acactaggaagcccatgtact-3'	22	189	278
Probe	TET-5'-tgtcctttgcagattcttgccttttca-3'-TAMRA	26	226	279
Reverse	5'-aattaatctagggctgtggaa-3'	22	254	280

CNS_neurodegeneration_v1.0 Summary: Ag3637 Expression of the CG59955-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

General_screening_panel_v1.4 Summary: Ag3637 Expression of the CG59955-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

Panel 4.1D Summary: Ag3637 Expression of the CG59955-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

General oncology screening panel_v_2.4 Summary: Ag3637 Expression of the CG59955-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

V. CG89211-01: GPCR

Expression of gene CG89211-01 was assessed using the primer-probe set Ag3691, described in Table VA. Results of the RTQ-PCR runs are shown in Table VB.

Table VA. Probe Name Ag3691

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - tgcattttaattcgctccagttc - 3'	22	464	281
Probe	TET-5' - cactcccgataatctatctcatctaccg - 3' - TAMRA	28	492	282
Reverse	5' - atgagcctgacaaaatggtaaa - 3'	22	520	283

5

Table VB. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3691, Run 267739062	Tissue Name	Rel. Exp.(%) Ag3691, Run 267739062
Colon cancer 1	0.0	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	0.0
Colon cancer 2	0.0	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	100.0
Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	0.0
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	67.4
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	0.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	0.0	Kidney NAT 1	0.0
Melanoma 2	0.0	Kidney cancer 2	0.0
Melanoma 3	0.0	Kidney NAT 2	0.0

metastatic melanoma 4	0.0	Kidney cancer 3	0.0
metastatic melanoma 5	88.3	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag3691 Expression of the CG89211-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

General_screening_panel_v1.4 Summary: Ag3691 Expression of the CG89211-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

5 **Panel 4.1D Summary:** Ag3691 Expression of the CG89211-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

General oncology screening panel_v_2.4 Summary: Ag3691 Expression of the CG89211-01 gene is low/undetectable in all samples on this panel (CTs>35). (Data not shown.)

10

W. CG90530-02: Ubiquitin-Conugating Enzyme

Expression of full length physical clone CG90530-02 was assessed using the primer-probe sets Ag3410 and Ag4036, described in Tables WA and WB. Results of the RTQ-PCR runs are shown in Tables WC, WD, WE and WF.

15

Table WA. Probe Name Ag3410

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tgagaggtacccatttgaacct-3'	22	372	284
Probe	TET-5'-cctcagatccgatttctcactccaat-3'-TAMRA	26	346	285
Reverse	5'-aaatccttccagcagaatcaat-3'	22	311	286

Table WB. Probe Name Ag4036

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tgagaggtacccatttgaacct-3'	22	372	287
Probe	TET-5'-cctcagatccgatttctcactccaat-3'-TAMRA	26	346	288
Reverse	5'-aaatccttccagcagaatcaat-3'	22	311	289

Table WC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3410, Run 210350165	Tissue Name	Rel. Exp.(%) Ag3410, Run 210350165
AD 1 Hippo	4.7	Control (Path) 3 Temporal Ctx	1.0
AD 2 Hippo	17.4	Control (Path) 4 Temporal Ctx	28.7
AD 3 Hippo	2.1	AD 1 Occipital Ctx	5.4
AD 4 Hippo	2.2	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	80.7	AD 3 Occipital Ctx	1.6
AD 6 Hippo	32.3	AD 4 Occipital Ctx	14.1
Control 2 Hippo	30.4	AD 5 Occipital Ctx	48.3
Control 4 Hippo	0.8	AD 6 Occipital Ctx	10.2
Control (Path) 3 Hippo	0.9	Control 1 Occipital Ctx	0.5
AD 1 Temporal Ctx	4.5	Control 2 Occipital Ctx	76.8
AD 2 Temporal Ctx	29.3	Control 3 Occipital Ctx	7.0
AD 3 Temporal Ctx	1.7	Control 4 Occipital Ctx	1.9
AD 4 Temporal Ctx	8.9	Control (Path) 1 Occipital Ctx	84.1
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	4.5
AD 5 Sup Temporal Ctx	20.7	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	29.3	Control (Path) 4 Occipital Ctx	7.7
AD 6 Sup Temporal Ctx	30.8	Control 1 Parietal Ctx	2.0
Control 1 Temporal Ctx	0.5	Control 2 Parietal Ctx	23.2
Control 2 Temporal Ctx	31.9	Control 3 Parietal Ctx	6.8
Control 3 Temporal Ctx	8.1	Control (Path) 1 Parietal Ctx	73.2
Control 3 Temporal Ctx	1.4	Control (Path) 2 Parietal Ctx	14.5
Control (Path) 1 Temporal Ctx	56.3	Control (Path) 3 Parietal Ctx	1.7
Control (Path) 2	23.7	Control (Path) 4 Parietal	26.6

Temporal Ctx		Ctx	
--------------	--	-----	--

Table WD. General screening panel_v1.4

Tissue Name	Rel. Exp.(%) Ag3410, Run 216839094	Rel. Exp.(%) Ag4036, Run 220283343	Tissue Name	Rel. Exp.(%) Ag3410, Run 216839094	Rel. Exp.(%) Ag4036, Run 220283343
Adipose	0.1	0.3	Renal ca. TK-10	14.0	21.9
Melanoma* Hs688(A).T	2.0	2.5	Bladder	2.2	1.7
Melanoma* Hs688(B).T	2.2	1.9	Gastric ca. (liver met.) NCI-N87	13.1	13.2
Melanoma* M14	45.1	34.9	Gastric ca. KATO III	85.3	90.1
Melanoma* LOXIMVI	11.0	17.4	Colon ca. SW-948	13.6	17.2
Melanoma* SK-MEL-5	55.5	44.8	Colon ca. SW480	100.0	100.0
Squamous cell carcinoma SCC-4	14.2	22.1	Colon ca.* (SW480 met) SW620	57.0	49.7
Testis Pool	2.3	8.0	Colon ca. HT29	21.9	22.8
Prostate ca.* (bone met) PC-3	3.6	5.7	Colon ca. HCT-116	82.9	86.5
Prostate Pool	0.3	0.0	Colon ca. CaCo-2	17.8	21.0
Placenta	0.7	0.8	Colon cancer tissue	4.2	3.6
Uterus Pool	0.2	0.0	Colon ca. SW1116	4.4	4.3
Ovarian ca. OVCAR-3	30.8	21.5	Colon ca. Colo-205	12.9	12.5
Ovarian ca. SK-OV-3	33.7	31.4	Colon ca. SW-48	14.9	26.2
Ovarian ca. OVCAR-4	27.2	30.6	Colon Pool	0.9	0.8
Ovarian ca. OVCAR-5	28.3	24.7	Small Intestine Pool	0.3	0.2
Ovarian ca. IGROV-1	9.4	9.5	Stomach Pool	0.4	0.3
Ovarian ca. OVCAR-8	8.2	6.2	Bone Marrow Pool	0.1	0.1
Ovary	0.6	0.6	Fetal Heart	5.3	5.0
Breast ca. MCF-7	32.1	25.2	Heart Pool	0.5	0.5
Breast ca. MDA-MB-231	47.0	51.8	Lymph Node Pool	2.1	1.5
Breast ca. BT 549	55.1	50.7	Fetal Skeletal Muscle	1.1	1.2
Breast ca. T47D	51.8	33.9	Skeletal Muscle Pool	2.2	1.5
Breast ca. MDA-N	24.5	15.9	Spleen Pool	0.7	0.6
Breast Pool	0.8	0.6	Thymus Pool	2.9	3.8
Trachea	0.5	0.5	CNS cancer	27.9	17.8

			(glio/astro) U87-MG		
Lung	0.1	0.0	CNS cancer (glio/astro) U-118-MG	49.0	48.6
Fetal Lung	3.1	2.2	CNS cancer (neuro;met) SK-N-AS	41.8	47.0
Lung ca. NCI-N417	27.9	25.0	CNS cancer (astro) SF-539	28.9	6.8
Lung ca. LX-1	40.9	38.7	CNS cancer (astro) SNB-75	34.4	29.7
Lung ca. NCI-H146	4.4	6.5	CNS cancer (glio) SNB-19	9.1	7.6
Lung ca. SHP-77	59.5	11.4	CNS cancer (glio) SF-295	4.6	5.4
Lung ca. A549	17.6	26.1	Brain (Amygdala) Pool	1.6	1.2
Lung ca. NCI-H526	11.6	12.0	Brain (cerebellum)	1.5	0.9
Lung ca. NCI-H23	14.6	22.1	Brain (fetal)	3.6	3.6
Lung ca. NCI-H460	2.5	2.4	Brain (Hippocampus) Pool	1.3	1.2
Lung ca. HOP-62	3.3	4.1	Cerebral Cortex Pool	2.4	1.4
Lung ca. NCI-H522	17.4	23.5	Brain (Substantia nigra) Pool	1.8	1.0
Liver	0.1	0.1	Brain (Thalamus) Pool	2.4	2.0
Fetal Liver	16.0	12.2	Brain (whole)	2.8	2.4
Liver ca. HepG2	4.0	6.1	Spinal Cord Pool	0.5	0.4
Kidney Pool	0.5	0.3	Adrenal Gland	0.2	0.4
Fetal Kidney	5.6	4.0	Pituitary gland Pool	0.1	0.3
Renal ca. 786-0	16.5	21.0	Salivary Gland	0.1	0.1
Renal ca. A498	3.1	2.8	Thyroid (female)	0.4	0.2
Renal ca. ACHN	10.2	10.8	Pancreatic ca. CAPAN2	52.5	39.2
Renal ca. UO-31	4.7	7.6	Pancreas Pool	1.1	1.0

Table WE. Panel 4D

Tissue Name	Rel. Exp.(%) Ag3410, Run 165296375	Tissue Name	Rel. Exp.(%) Ag3410, Run 165296375
Secondary Th1 act	29.5	HUVEC IL-1beta	3.4
Secondary Th2 act	19.5	HUVEC IFN gamma	6.2

Secondary Tr1 act	29.9	HUVEC TNF alpha + IFN gamma	5.1
Secondary Th1 rest	0.8	HUVEC TNF alpha + IL4	4.6
Secondary Th2 rest	1.3	HUVEC IL-11	4.0
Secondary Tr1 rest	1.1	Lung Microvascular EC none	3.8
Primary Th1 act	10.6	Lung Microvascular EC TNFalpha + IL-1beta	3.8
Primary Th2 act	11.7	Microvascular Dermal EC none	6.0
Primary Tr1 act	25.7	Microvascular Dermal EC TNFalpha + IL-1beta	3.7
Primary Th1 rest	23.7	Bronchial epithelium TNFalpha + IL1beta	1.6
Primary Th2 rest	4.5	Small airway epithelium none	0.3
Primary Tr1 rest	20.3	Small airway epithelium TNFalpha + IL-1beta	5.1
CD45RA CD4 lymphocyte act	13.1	Coronary artery SMC rest	2.5
CD45RO CD4 lymphocyte act	16.0	Coronary artery SMC TNFalpha + IL-1beta	3.2
CD8 lymphocyte act	13.9	Astrocytes rest	1.3
Secondary CD8 lymphocyte rest	20.2	Astrocytes TNFalpha + IL-1beta	0.6
Secondary CD8 lymphocyte act	13.7	KU-812 (Basophil) rest	10.9
CD4 lymphocyte none	0.1	KU-812 (Basophil) PMA/ionomycin	21.8
2ry Th1/Th2/Tr1_anti-CD95 CH11	3.0	CCD1106 (Keratinocytes) none	12.4
LAK cells rest	2.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	2.6
LAK cells IL-2	9.5	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	19.9	Lupus kidney	0.1
LAK cells IL-2+IFN gamma	23.0	NCI-H292 none	12.2
LAK cells IL-2+ IL-18	17.3	NCI-H292 IL-4	30.8
LAK cells PMA/ionomycin	0.8	NCI-H292 IL-9	25.7
NK Cells IL-2 rest	10.6	NCI-H292 IL-13	9.9
Two Way MLR 3 day	3.0	NCI-H292 IFN gamma	12.4

Two Way MLR 5 day	9.0	HPAEC none	4.5
Two Way MLR 7 day	7.1	HPAEC TNF alpha + IL-1 beta	2.5
PBMC rest	0.2	Lung fibroblast none	1.7
PBMC PWM	47.0	Lung fibroblast TNF alpha + IL-1 beta	2.1
PBMC PHA-L	18.2	Lung fibroblast IL-4	2.3
Ramos (B cell) none	23.7	Lung fibroblast IL-9	3.4
Ramos (B cell) ionomycin	39.5	Lung fibroblast IL-13	1.9
B lymphocytes PWM	100.0	Lung fibroblast IFN gamma	1.2
B lymphocytes CD40L and IL-4	9.3	Dermal fibroblast CCD1070 rest	21.0
EOL-1 dbcAMP	4.8	Dermal fibroblast CCD1070 TNF alpha	51.4
EOL-1 dbcAMP PMA/ionomycin	4.7	Dermal fibroblast CCD1070 IL-1 beta	9.8
Dendritic cells none	0.8	Dermal fibroblast IFN gamma	7.0
Dendritic cells LPS	0.1	Dermal fibroblast IL-4	5.8
Dendritic cells anti-CD40	0.3	IBD Colitis 2	0.4
Monocytes rest	0.1	IBD Crohn's	0.1
Monocytes LPS	0.0	Colon	0.9
Macrophages rest	2.7	Lung	0.6
Macrophages LPS	0.1	Thymus	0.4
HUVEC none	8.0	Kidney	8.1
HUVEC starved	28.1		

Table WF. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3410, Run 267171484	Rel. Exp.(%) Ag4036, Run 268362917	Tissue Name	Rel. Exp.(%) Ag3410, Run 267171484	Rel. Exp.(%) Ag4036, Run 268362917
Colon cancer 1	34.4	11.3	Bladder cancer NAT 2	0.1	0.0
Colon cancer NAT 1	12.3	10.8	Bladder cancer NAT 3	0.1	0.0
Colon cancer 2	42.9	7.1	Bladder cancer NAT 4	0.6	0.0
Colon cancer NAT	2.9	7.0	Adenocarcinoma of the prostate 1	3.2	0.0

2					
Colon cancer 3	70.2	47.3	Adenocarcinoma of the prostate 2	0.7	0.0
Colon cancer NAT 3	14.6	18.0	Adenocarcinoma of the prostate 3	0.9	0.0
Colon malignant cancer 4	90.8	68.8	Adenocarcinoma of the prostate 4	11.7	6.6
Colon normal adjacent tissue 4	6.2	1.7	Prostate cancer NAT 5	1.3	0.6
Lung cancer 1	10.2	0.8	Adenocarcinoma of the prostate 6	0.5	0.0
Lung NAT 1	0.1	0.0	Adenocarcinoma of the prostate 7	0.8	0.0
Lung cancer 2	100.0	100.0	Adenocarcinoma of the prostate 8	0.1	0.0
Lung NAT 2	0.1	0.0	Adenocarcinoma of the prostate 9	5.3	1.7
Squamous cell carcinoma 3	26.2	17.6	Prostate cancer NAT 10	0.1	0.0
Lung NAT 3	0.2	0.0	Kidney cancer 1	4.6	3.3
metastatic melanoma 1	0.9	0.0	Kidney NAT 1	0.6	0.5
Melanoma 2	0.2	0.0	Kidney cancer 2	21.0	15.0
Melanoma 3	0.9	0.0	Kidney NAT 2	2.9	2.0
metastatic melanoma 4	7.6	3.4	Kidney cancer 3	2.8	1.5
metastatic melanoma 5	7.7	3.1	Kidney NAT 3	0.7	0.7
Bladder cancer 1	0.3	0.0	Kidney cancer 4	2.1	2.0
Bladder cancer NAT 1	0.0	0.0	Kidney NAT 4	0.6	1.0
Bladder cancer 2	0.4	0.2			

CNS_neurodegeneration_v1.0 Summary: Ag3410 This panel does not show differential expression of the CG90530-02-01 gene in Alzheimer's disease. However, this

expression profile confirms the presence of this gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system.

General_screening_panel_v1.4 Summary: Ag3410/Ag4036 Two experiments with the same probe and primer set produce results that are in excellent agreement.

5 Highest expression of the CG90530-02 gene is seen in a colon cancer cell line (CTs=25-26). In addition, expression is significantly higher in all the cancer cell lines on this panel when compared to expression in the normal tissues. Thus, expression of this gene could be used as a marker of cancer. In addition, this gene is expressed at much higher levels in fetal lung, liver, kidney and heart tissue (CTs=28-30) when compared to expression in the adult counterpart (CTs=33-36). Thus, expression of this gene may be used to differentiate between the fetal and adult source of these tissues. This gene encodes a putative member of the ubiquitin conjugating enzyme family. Ubiquitin-dependent protein degradation plays a role in many cellular processes and has been shown to be upregulated in some cancers (Eliseeva E. Cell Growth Differ 2001 Aug;12(8):427-33) Furthermore, higher levels of expression of this gene in cancer cell lines and fetal tissues suggests that therapeutic modulation of the expression or function of this gene may be useful in the treatment of cancer.

Among tissues with metabolic function, this gene is expressed at moderate to low levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that 20 disregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

This gene is also expressed at moderate levels in the CNS, including the 25 hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 4D Summary: Ag3410 Highest expression of the CG90530-02 is seen in 30 PWM stimulated B lymphocytes (CT=23.7). High levels of expression are also seen in activated T cells, LAK cells, dermal fibroblasts and the pulmonary mucoepidermoid cell line NCI-H292. Thus, therapeutic regulation of the transcript or the protein encoded by the

transcript could be important in immune modulation and in the treatment of T and B cell-mediated diseases such as asthma, arthritis, psoriasis, IBD, and lupus.

General oncology screening panel_v_2.4 Summary: Ag3410/Ag4036 Two experiments with the same probe and primer set produce results that are in excellent agreement. Highest expression of the CG90530-02 gene is seen in lung cancer (CTs=28-31). In addition, this gene appears to be overexpressed in lung, kidney and colon cancers when compared to expression in normal adjacent tissue. Thus, expression of this gene could be used as a marker of lung and colon cancer. Ubiquitinylation is a cyclical process operating in all cells to target specific proteins (eg, p53) for degradation. Abnormal accumulations of ubiquitinylated proteins have been identified in colorectal carcinoma. This gene encodes a putative ubiquitin conjugating enzyme. Therefore, therapeutic modulation of the expression or function of this gene could be effective in the treatment of lung, kidney and colon cancer.

15 X. CG93076-01: GPCR

Expression of full length physical clone CG93076-01 was assessed using the primer-probe set Ag2129, described in Table XA. Results of the RTQ-PCR runs are shown in Tables XB and XC.

Table XA. Probe Name Ag2129

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gcaacagcatggtgatctg-3'	19	133	290
Probe	TET-5'-ctttcgaatgcacaggaaccccttct-3'-TAMRA	26	161	291
Reverse	5'-cgccaggttgaggatatagat-3'	21	189	292

20

Table XB. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag2129, Run 148719315	Tissue Name	Rel. Exp.(%) Ag2129, Run 148719315
Liver adenocarcinoma	11.7	Kidney (fetal)	0.0
Pancreas	0.0	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	22.8
Adrenal gland	6.8	Renal ca. RXF 393	0.0
Thyroid	0.0	Renal ca. ACHN	0.0
Salivary gland	0.0	Renal ca. UO-31	0.0

Pituitary gland	11.8	Renal ca. TK-10	11.4
Brain (fetal)	51.8	Liver	0.0
Brain (whole)	14.9	Liver (fetal)	24.0
Brain (amygdala)	14.2	Liver ca. (hepatoblast) HepG2	9.5
Brain (cerebellum)	55.1	Lung	0.0
Brain (hippocampus)	11.1	Lung (fetal)	24.3
Brain (substantia nigra)	0.0	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	12.2	Lung ca. (small cell) NCI-H69	0.0
Cerebral Cortex	0.0	Lung ca. (s.cell var.) SHP-77	0.0
Spinal cord	0.0	Lung ca. (large cell) NCI-H460	0.0
glio/astro U87-MG	0.0	Lung ca. (non-sm. cell) A549	24.8
glio/astro U-118- MG	0.0	Lung ca. (non-s.cell) NCI-H23	100.0
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP-62	0.0
neuro*; met SK-N- AS	5.5	Lung ca. (non-s.cl) NCI-H522	9.2
astrocytoma SF-539	0.0	Lung ca. (squam.) SW 900	0.0
astrocytoma SNB- 75	0.0	Lung ca. (squam.) NCI-H596	0.0
glioma SNB-19	0.0	Mammary gland	4.3
glioma U251	21.6	Breast ca.* (pl.ef) MCF-7	0.0
glioma SF-295	8.7	Breast ca.* (pl.ef) MDA-MB-231	0.0
Heart (fetal)	0.0	Breast ca.* (pl.ef) T47D	0.0
Heart	0.0	Breast ca. BT-549	0.0
Skeletal muscle (fetal)	78.5	Breast ca. MDA-N	9.8
Skeletal muscle	0.0	Ovary	0.0
Bone marrow	12.5	Ovarian ca. OVCAR-3	10.4
Thymus	14.9	Ovarian ca. OVCAR-4	0.0
Spleen	0.0	Ovarian ca. OVCAR-5	12.2

Lymph node	28.7	Ovarian ca. OVCAR-8	16.5
Colorectal	76.3	Ovarian ca. IGROV-1	22.1
Stomach	0.0	Ovarian ca.* (ascites) SK-OV-3	0.0
Small intestine	0.0	Uterus	81.8
Colon ca. SW480	19.2	Placenta	0.0
Colon ca.* SW620(SW480 met)	0.0	Prostate	0.0
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	0.0
Colon ca. HCT-116	0.0	Testis	55.1
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.0
Colon ca. tissue(ODO3866)	0.0	Melanoma* (met) Hs688(B).T	0.0
Colon ca. HCC- 2998	0.0	Melanoma UACC-62	0.0
Gastric ca.* (liver met) NCI-N87	14.5	Melanoma M14	0.0
Bladder	0.0	Melanoma LOX IMVI	0.0
Trachea	0.0	Melanoma* (met) SK- MEL-5	16.5
Kidney	0.0	Adipose	20.3

Table XC. Panel 4D

Tissue Name	Rel. Exp.(%) Ag2129, Run 165823129	Tissue Name	Rel. Exp.(%) Ag2129, Run 165823129
Secondary Th1 act	3.0	HUVEC IL-1beta	0.0
Secondary Th2 act	12.6	HUVEC IFN gamma	3.8
Secondary Tr1 act	7.5	HUVEC TNF alpha + IFN gamma	2.6
Secondary Th1 rest	12.3	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	8.7	HUVEC IL-11	0.0
Secondary Tr1 rest	15.7	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	2.3	Microvascular Dermal EC none	0.0
Primary Tr1 act	5.2	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	84.7	Bronchial epithelium TNFalpha + IL1beta	4.8

Primary Th2 rest	34.6	Small airway epithelium none	0.0
Primary Tr1 rest	17.1	Small airway epithelium TNFalpha + IL-1beta	5.7
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	4.3
CD45RO CD4 lymphocyte act	6.5	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	2.4	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	9.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	1.3	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	39.8	KU-812 (Basophil) PMA/ionomycin	5.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	45.7	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	9.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	31.9	Liver cirrhosis	100.0
LAK cells IL-2+IL-12	17.6	Lupus kidney	0.0
LAK cells IL-2+IFN gamma	20.0	NCI-H292 none	7.3
LAK cells IL-2+ IL-18	13.3	NCI-H292 IL-4	10.6
LAK cells PMA/ionomycin	10.1	NCI-H292 IL-9	4.1
NK Cells IL-2 rest	8.5	NCI-H292 IL-13	6.3
Two Way MLR 3 day	9.2	NCI-H292 IFN gamma	1.2
Two Way MLR 5 day	2.5	HPAEC none	0.0
Two Way MLR 7 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	8.1	Lung fibroblast none	2.5
PBMC PWM	3.8	Lung fibroblast TNF alpha + IL-1 beta	9.0
PBMC PHA-L	1.7	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	0.0
B lymphocytes	5.8	Lung fibroblast IFN gamma	0.0

PWM			
B lymphocytes CD40L and IL-4	14.2	Dermal fibroblast CCD1070 rest	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 TNF alpha	5.7
EOL-1 dbcAMP PMA/ionomycin	8.7	Dermal fibroblast CCD1070 IL-1 beta	0.0
Dendritic cells none	6.6	Dermal fibroblast IFN gamma	0.0
Dendritic cells LPS	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells anti- CD40	5.1	IBD Colitis 2	5.3
Monocytes rest	3.4	IBD Crohn's	4.7
Monocytes LPS	2.5	Colon	97.9
Macrophages rest	0.0	Lung	28.7
Macrophages LPS	0.0	Thymus	2.5
HUVEC none	0.0	Kidney	21.2
HUVEC starved	0.0		

Panel 1.3D Summary: Ag2129 Expression of the CG93076-01 gene is restricted to a sample derived from a lung cancer cell line (CT=34.3). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of lung cancer. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of lung cancer.

Panel 4D Summary: Ag2129 Expression of the CG93076-01 gene is highest in liver cirrhosis (CT=31.8). Low but significant levels of expression are seen in primary resting T cells, colon, kidney, lung, LAK cells, and untreated CD4s. Expression of this gene is not detected in normal liver in Panel 1.3D, suggesting that its expression is unique to liver cirrhosis. This gene encodes a putative GPCR; therefore, antibodies or small molecule therapeutics could reduce or inhibit fibrosis that occurs in liver cirrhosis. In addition, antibodies to this putative GPCR could also be used for the diagnosis of liver cirrhosis.

15 Y. CG94235-01: Thymidylate Kinase

Expression of gene CG94235-01 was assessed using the primer-probe sets Ag1980 and Ag3909, described in Tables YA and YB. Results of the RTQ-PCR runs are shown in Tables YC, YD, YE, YF, YG, YH, YI and YJ.

Table YA. Probe Name Ag1980

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gggtacgatggctgaagtaaa-3'	21	1437	293
Probe	TET-5'-ccagttttctgccacacaacatgctt-3'-TAMRA	26	1394	294
Reverse	5'-ttatgcagtgttcccaaatttc-3'	22	1364	295

Table YB. Probe Name Ag3909

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-caggtgccacgtctaactagat-3'	22	1307	296
Probe	TET-5'-tggtgtttgaaacatctacatccacca-3'-TAMRA	27	1333	297
Reverse	5'-gaaatttggaacactgcataa-3'	22	1364	298

Table YC. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag1980. Run 211061884	Rel. Exp.(%) Ag1980. Run 212317511	Tissue Name	Rel. Exp.(%) Ag1980. Run 211061884	Rel. Exp.(%) Ag1980. Run 212317511
110967 COPD-F	10.1	11.6	112427 Match Control Psoriasis-F	100.0	100.0
110980 COPD-F	12.5	16.8	112418 Psoriasis-M	19.1	18.7
110968 COPD-M	13.9	13.5	112723 Match Control Psoriasis-M	6.6	7.2
110977 COPD-M	50.3	32.8	112419 Psoriasis-M	35.6	17.0
110989 Emphysema-F	20.7	24.5	112424 Match Control Psoriasis-M	7.6	10.1
110992 Emphysema-F	12.2	6.0	112420 Psoriasis-M	54.0	53.2
110993 Emphysema-F	12.9	8.7	112425 Match Control Psoriasis-M	62.4	72.7
110994 Emphysema-F	7.0	5.8	104689 (MF) OA Bone-Backus	66.0	56.3
110995 Emphysema-F	17.0	17.6	104690 (MF) Adj "Normal" Bone-Backus	34.6	24.0
110996 Emphysema-F	3.3	7.0	104691 (MF) OA Synovium-Backus	36.6	40.9
110997 Asthma-M	8.0	3.1	104692 (BA) OA Cartilage-Backus	12.1	6.8
111001 Asthma-F	21.8	15.8	104694 (BA) OA Bone-Backus	51.8	36.3
111002	24.0	13.1	104695 (BA) Adj	26.4	19.5

Asthma-F			"Normal" Bone-Backus		
111003 Atopic Asthma-F	20.3	22.5	104696 (BA) OA Synovium-Backus	63.7	40.9
111004 Atopic Asthma-F	26.1	28.5	104700 (SS) OA Bone-Backus	20.9	62.0
111005 Atopic Asthma-F	24.5	10.3	104701 (SS) Adj "Normal" Bone-Backus	28.7	25.7
111006 Atopic Asthma-F	11.0	8.1	104702 (SS) OA Synovium-Backus	80.1	59.5
111417 Allergy-M	8.7	7.7	117093 OA Cartilage Rep7	17.8	20.9
112347 Allergy-M	0.0	0.0	112672 OA Bone5	36.6	33.9
112349 Normal Lung-F	0.0	0.0	112673 OA Synovium5	21.3	22.4
112357 Normal Lung-F	67.4	64.6	112674 OA Synovial Fluid cells5	20.3	20.0
112354 Normal Lung-M	9.8	15.6	117100 OA Cartilage Rep14	9.6	6.9
112374 Crohns-F	11.1	15.0	112756 OA Bone9	95.3	66.0
112389 Match Control Crohns-F	10.4	18.4	112757 OA Synovium9	17.4	19.2
112375 Crohns-F	14.5	9.9	112758 OA Synovial Fluid Cells9	14.1	17.8
112732 Match Control Crohns-F	29.3	28.5	117125 RA Cartilage Rep2	19.9	22.5
112725 Crohns-M	5.1	3.6	113492 Bone2 RA	76.3	66.0
112387 Match Control Crohns-M	14.4	15.7	113493 Synovium2 RA	20.9	19.3
112378 Crohns-M	0.0	0.0	113494 Syn Fluid Cells RA	48.0	43.2
112390 Match Control Crohns-M	33.7	48.6	113499 Cartilage4 RA	40.3	49.7
112726 Crohns-M	25.0	19.9	113500 Bone4 RA	63.3	50.3
112731 Match Control	19.1	16.8	113501 Synovium4 RA	44.8	36.6

Crohns-M					
112380 Ulcer Col-F	14.8	15.9	113502 Syn Fluid Cells4 RA	33.7	21.5
112734 Match Control Ulcer Col-F	74.2	59.9	113495 Cartilage3 RA	48.3	29.3
112384 Ulcer Col-F	26.2	28.9	113496 Bone3 RA	51.8	45.1
112737 Match Control Ulcer Col-F	8.8	4.1	113497 Synovium3 RA	21.9	25.3
112386 Ulcer Col-F	11.8	10.9	113498 Syn Fluid Cells3 RA	44.1	47.6
112738 Match Control Ulcer Col-F	27.0	15.0	117106 Normal Cartilage Rep20	2.0	5.3
112381 Ulcer Col-M	0.0	0.0	113663 Bone3 Normal	0.0	0.2
112735 Match Control Ulcer Col-M	6.3	6.5	113664 Synovium3 Normal	0.0	0.0
112382 Ulcer Col-M	13.5	10.8	113665 Syn Fluid Cells3 Normal	0.0	0.0
112394 Match Control Ulcer Col-M	1.6	7.7	117107 Normal Cartilage Rep22	10.6	11.8
112383 Ulcer Col-M	18.7	26.6	113667 Bone4 Normal	10.1	8.4
112736 Match Control Ulcer Col-M	12.9	6.7	113668 Synovium4 Normal	10.6	8.1
112423 Psoriasis-F	20.7	18.6	113669 Syn Fluid Cells4 Normal	15.8	21.3

Table YD. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3909, Run 212248424	Tissue Name	Rel. Exp.(%) Ag3909, Run 212248424
AD 1 Hippo	11.7	Control (Path) 3 Temporal Ctx	5.6
AD 2 Hippo	41.5	Control (Path) 4 Temporal Ctx	33.7
AD 3 Hippo	7.6	AD 1 Occipital Ctx	12.5
AD 4 Hippo	7.2	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	75.3	AD 3 Occipital Ctx	1.1
AD 6 Hippo	36.1	AD 4 Occipital Ctx	23.7

Control 2 Hippo	19.1	AD 5 Occipital Ctx	29.9
Control 4 Hippo	12.9	AD 6 Occipital Ctx	35.4
Control (Path) 3 Hippo	12.2	Control 1 Occipital Ctx	9.1
AD 1 Temporal Ctx	9.5	Control 2 Occipital Ctx	69.7
AD 2 Temporal Ctx	36.3	Control 3 Occipital Ctx	18.9
AD 3 Temporal Ctx	3.7	Control 4 Occipital Ctx	3.9
AD 4 Temporal Ctx	31.9	Control (Path) 1 Occipital Ctx	71.2
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	31.9
AD 5 Sup Temporal Ctx	56.6	Control (Path) 3 Occipital Ctx	4.3
AD 6 Inf Temporal Ctx	22.4	Control (Path) 4 Occipital Ctx	23.0
AD 6 Sup Temporal Ctx	32.3	Control 1 Parietal Ctx	8.5
Control 1 Temporal Ctx	5.3	Control 2 Parietal Ctx	38.2
Control 2 Temporal Ctx	42.9	Control 3 Parietal Ctx	29.1
Control 3 Temporal Ctx	13.0	Control (Path) 1 Parietal Ctx	73.2
Control 4 Temporal Ctx	4.4	Control (Path) 2 Parietal Ctx	44.4
Control (Path) 1 Temporal Ctx	58.2	Control (Path) 3 Parietal Ctx	7.3
Control (Path) 2 Temporal Ctx	41.8	Control (Path) 4 Parietal Ctx	49.0

Table YE. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag3909, Run 2172358 26	Rel. Exp.(%) Ag3909, Run 219173644	Tissue Name	Rel. Exp.(%) Ag3909, Run 217235826	Rel. Exp.(%) Ag3909, Run 2191736 44
Adipose	0.6	0.6	Renal ca. TK-10	0.0	0.0
Melanoma* Hs688(A).T	0.0	0.0	Bladder	9.7	11.7
Melanoma* Hs688(B).T	0.0	0.0	Gastric ca. (liver met.) NCI-N87	100.0	100.0

Melanoma* M14	0.3	0.4	Gastric ca. KATO III	0.5	0.6
Melanoma* LOXIMVI	0.1	0.2	Colon ca. SW-948	0.5	0.5
Melanoma* SK-MEL-5	0.2	0.2	Colon ca. SW480	0.1	0.1
Squamous cell carcinoma SCC-4	0.5	0.5	Colon ca.* (SW480 met) SW620	0.0	0.0
Testis Pool	0.6	0.5	Colon ca. HT29	0.1	0.1
Prostate ca.* (bone met) PC-3	0.1	0.1	Colon ca. HCT-116	0.3	0.3
Prostate Pool	0.3	0.3	Colon ca. CaCo-2	0.1	0.0
Placenta	0.4	0.5	Colon cancer tissue	0.6	0.5
Uterus Pool	0.2	0.2	Colon ca. SW1116	0.2	0.1
Ovarian ca. OVCAR-3	0.2	0.2	Colon ca. Colo-205	0.0	0.0
Ovarian ca. SK-OV-3	0.5	0.4	Colon ca. SW-48	0.0	0.0
Ovarian ca. OVCAR-4	0.1	0.1	Colon Pool	0.4	0.5
Ovarian ca. OVCAR-5	0.6	0.5	Small Intestine Pool	0.4	0.4
Ovarian ca. IGROV-1	0.0	0.0	Stomach Pool	0.1	0.3
Ovarian ca. OVCAR-8	0.0	0.0	Bone Marrow Pool	0.2	0.2
Ovary	0.8	0.6	Fetal Heart	0.2	0.2
Breast ca. MCF-7	0.0	0.0	Heart Pool	0.5	0.5
Breast ca. MDA-MB-231	0.2	0.2	Lymph Node Pool	0.4	0.4
Breast ca. BT 549	2.3	3.1	Fetal Skeletal Muscle	0.2	0.2
Breast ca. T47D	1.1	1.0	Skeletal Muscle Pool	1.7	1.9
Breast ca. MDA-N	0.2	0.2	Spleen Pool	2.3	2.6
Breast Pool	0.3	0.4	Thymus Pool	0.5	0.5
Trachea	0.9	1.0	CNS cancer (glio/astro) U87-MG	0.0	0.0
Lung	0.1	0.1	CNS cancer (glio/astro) U-118-MG	1.0	1.3
Fetal Lung	0.9	0.9	CNS cancer (neuro;met) SK-N-AS	1.2	1.5
Lung ca. NCI-N417	0.0	0.0	CNS cancer (astro) SF- 539	0.3	0.4
Lung ca. LX-1	0.0	0.0	CNS cancer (astro) SNB-75	0.1	0.1
Lung ca. NCI-H146	0.9	1.1	CNS cancer (glio) SNB-19	0.0	0.0
Lung ca. SHP-77	0.5	0.5	CNS cancer (glio) SF- 295	0.2	0.2
Lung ca. A549	0.0	0.0	Brain (Amygdala) Pool	0.6	0.5
Lung ca. NCI-H526	0.0	0.0	Brain (cerebellum)	0.1	0.2

Lung ca. NCI-H23	0.0	0.0	Brain (fetal)	0.8	0.7
Lung ca. NCI-H460	0.0	0.0	Brain (Hippocampus) Pool	0.5	0.5
Lung ca. HOP-62	0.0	0.0	Cerebral Cortex Pool	0.7	0.7
Lung ca. NCI-H522	0.0	0.0	Brain (Substantia nigra) Pool	0.8	0.7
Liver	0.1	0.1	Brain (Thalamus) Pool	0.8	0.9
Fetal Liver	4.7	4.1	Brain (whole)	1.3	1.3
Liver ca. HepG2	0.0	0.0	Spinal Cord Pool	0.4	0.5
Kidney Pool	0.7	0.8	Adrenal Gland	0.7	0.8
Fetal Kidney	0.2	0.2	Pituitary gland Pool	0.3	0.3
Renal ca. 786-0	0.0	0.0	Salivary Gland	0.5	0.4
Renal ca. A498	0.2	0.2	Thyroid (female)	0.4	0.2
Renal ca. ACHN	0.0	0.0	Pancreatic ca. CAPAN2	0.3	0.3
Renal ca. UO-31	0.1	0.0	Pancreas Pool	0.4	0.3

Table YF. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag1980, Run 165534458	Tissue Name	Rel. Exp.(%) Ag1980, Run 165534458
Liver adenocarcinoma	1.4	Kidney (fetal)	0.4
Pancreas	0.4	Renal ca. 786-0	0.0
Pancreatic ca. CAPAN 2	0.0	Renal ca. A498	2.0
Adrenal gland	0.6	Renal ca. RXF 393	0.7
Thyroid	0.7	Renal ca. ACHN	0.0
Salivary gland	1.0	Renal ca. UO-31	0.0
Pituitary gland	0.5	Renal ca. TK-10	0.1
Brain (fetal)	0.8	Liver	0.5
Brain (whole)	3.8	Liver (fetal)	5.3
Brain (amygdala)	2.9	Liver ca. (hepatoblast) HepG2	0.0
Brain (cerebellum)	0.6	Lung	1.5
Brain (hippocampus)	3.2	Lung (fetal)	2.5
Brain (substantia nigra)	1.6	Lung ca. (small cell) LX-1	0.0
Brain (thalamus)	5.1	Lung ca. (small cell) NCI-H69	0.4
Cerebral Cortex	1.9	Lung ca. (s.cell var.) SHP-77	0.5

Spinal cord	1.6	Lung ca. (large cell)NCI-H460	0.0
glio/astro U87-MG	0.0	Lung ca. (non-sm. cell) A549	0.0
glio/astro U-118-MG	1.8	Lung ca. (non-s.cell) NCI-H23	0.0
astrocytoma SW1783	0.0	Lung ca. (non-s.cell) HOP-62	0.0
neuro*; met SK-N-AS	1.1	Lung ca. (non-s.cl) NCI-H522	0.0
astrocytoma SF-539	7.9	Lung ca. (squam.) SW 900	1.5
astrocytoma SNB-75	1.7	Lung ca. (squam.) NCI-H596	0.1
glioma SNB-19	0.0	Mammary gland	0.7
glioma U251	0.7	Breast ca.* (pl.ef) MCF-7	0.1
glioma SF-295	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.3
Heart (fetal)	0.1	Breast ca.* (pl.ef) T47D	0.0
Heart	1.2	Breast ca. BT-549	1.1
Skeletal muscle (fetal)	0.0	Breast ca. MDA-N	0.0
Skeletal muscle	1.0	Ovary	0.3
Bone marrow	4.5	Ovarian ca. OVCAR-3	0.1
Thymus	1.2	Ovarian ca. OVCAR-4	0.1
Spleen	1.4	Ovarian ca. OVCAR-5	0.3
Lymph node	2.9	Ovarian ca. OVCAR-8	0.0
Colorectal	0.3	Ovarian ca. IGROV-1	0.0
Stomach	1.4	Ovarian ca.* (ascites) SK-OV-3	0.4
Small intestine	1.2	Uterus	1.8
Colon ca. SW480	0.0	Placenta	1.3
Colon ca.* SW620(SW480 met)	0.0	Prostate	0.3
Colon ca. HT29	0.0	Prostate ca.* (bone met)PC-3	0.1
Colon ca. HCT-116	0.1	Testis	1.2
Colon ca. CaCo-2	0.0	Melanoma Hs688(A).T	0.1
Colon ca. tissue(ODO3866)	0.6	Melanoma* (met) Hs688(B).T	0.1
Colon ca. HCC-2998	12.0	Melanoma UACC-62	1.0
Gastric ca.* (liver	100.0	Melanoma M14	0.3

met) NCI-N87			
Bladder	8.2	Melanoma LOX IMVI	0.1
Trachea	2.0	Melanoma* (met) SK-MEL-5	0.1
Kidney	0.4	Adipose	0.6

Table YG. Panel 2D

Tissue Name	Rel. Exp.(%) Ag1980, Run 169484147	Tissue Name	Rel. Exp.(%) Ag1980, Run 169484147
Normal Colon	14.7	Kidney Margin 8120608	1.1
CC Well to Mod Diff (ODO3866)	3.2	Kidney Cancer 8120613	1.4
CC Margin (ODO3866)	6.3	Kidney Margin 8120614	1.5
CC Gr.2 rectosigmoid (ODO3868)	3.5	Kidney Cancer 9010320	6.2
CC Margin (ODO3868)	3.3	Kidney Margin 9010321	2.6
CC Mod Diff (ODO3920)	1.3	Normal Uterus	1.2
CC Margin (ODO3920)	1.5	Uterus Cancer 064011	3.4
CC Gr.2 ascend colon (ODO3921)	3.7	Normal Thyroid	4.7
CC Margin (ODO3921)	1.8	Thyroid Cancer 064010	12.1
CC from Partial Hepatectomy (ODO4309) Mets	5.0	Thyroid Cancer A302152	1.6
Liver Margin (ODO4309)	6.7	Thyroid Margin A302153	4.4
Colon mets to lung (OD04451-01)	5.0	Normal Breast	4.5
Lung Margin (OD04451-02)	8.4	Breast Cancer (OD04566)	8.1
Normal Prostate 6546-1	5.2	Breast Cancer (OD04590-01)	11.8
Prostate Cancer (OD04410)	9.1	Breast Cancer Mets (OD04590-03)	7.9
Prostate Margin (OD04410)	3.1	Breast Cancer Metastasis (OD04655-05)	5.3
Prostate Cancer (OD04720-01)	3.2	Breast Cancer 064006	18.0

Prostate Margin (OD04720-02)	6.4	Breast Cancer 1024	2.7
Normal Lung 061010	14.8	Breast Cancer 9100266	6.2
Lung Met to Muscle (ODO4286)	3.4	Breast Margin 9100265	1.9
Muscle Margin (ODO4286)	2.8	Breast Cancer A209073	6.5
Lung Malignant Cancer (OD03126)	5.2	Breast Margin A209073	1.5
Lung Margin (OD03126)	21.5	Normal Liver	2.3
Lung Cancer (OD04404)	53.6	Liver Cancer 064003	6.2
Lung Margin (OD04404)	6.8	Liver Cancer 1025	6.5
Lung Cancer (OD04565)	6.8	Liver Cancer 1026	2.1
Lung Margin (OD04565)	3.6	Liver Cancer 6004-T	6.0
Lung Cancer (OD04237-01)	10.1	Liver Tissue 6004-N	2.9
Lung Margin (OD04237-02)	24.5	Liver Cancer 6005-T	3.3
Ocular Mel Met to Liver (ODO4310)	0.2	Liver Tissue 6005-N	1.7
Liver Margin (ODO4310)	3.5	Normal Bladder	100.0
Melanoma Mets to Lung (OD04321)	1.8	Bladder Cancer 1023	0.6
Lung Margin (OD04321)	10.3	Bladder Cancer A302173	6.3
Normal Kidney	11.4	Bladder Cancer (OD04718-01)	74.2
Kidney Ca, Nuclear grade 2 (OD04338)	5.3	Bladder Normal Adjacent (OD04718-03)	5.1
Kidney Margin (OD04338)	4.2	Normal Ovary	3.5
Kidney Ca Nuclear grade 1/2 (OD04339)	1.5	Ovarian Cancer 064008	11.7
Kidney Margin (OD04339)	4.6	Ovarian Cancer (OD04768-07)	99.3
Kidney Ca, Clear cell type (OD04340)	11.1	Ovary Margin (OD04768-08)	0.9
Kidney Margin	3.5	Normal Stomach	10.7

(OD04340)			
Kidney Ca, Nuclear grade 3 (OD04348)	59.9	Gastric Cancer 9060358	3.0
Kidney Margin (OD04348)	81.2	Stomach Margin 9060359	9.9
Kidney Cancer (OD04622-01)	6.1	Gastric Cancer 9060395	6.9
Kidney Margin (OD04622-03)	1.1	Stomach Margin 9060394	7.7
Kidney Cancer (OD04450-01)	1.5	Gastric Cancer 9060397	7.2
Kidney Margin (OD04450-03)	1.7	Stomach Margin 9060396	10.4
Kidney Cancer 8120607	0.9	Gastric Cancer 064005	19.9

Table YH, Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3909, Run 170127176	Tissue Name	Rel. Exp.(%) Ag3909, Run 170127176
Secondary Th1 act	0.9	HUVEC IL-1beta	0.1
Secondary Th2 act	32.3	HUVEC IFN gamma	0.9
Secondary Tr1 act	3.6	HUVEC TNF alpha + IFN gamma	10.5
Secondary Th1 rest	5.5	HUVEC TNF alpha + IL4	0.4
Secondary Th2 rest	2.0	HUVEC IL-11	0.0
Secondary Tr1 rest	4.8	Lung Microvascular EC none	0.2
Primary Th1 act	1.3	Lung Microvascular EC TNFalpha + IL-1beta	0.6
Primary Th2 act	2.0	Microvascular Dermal EC none	0.1
Primary Tr1 act	1.1	Microvascular Dermal EC TNFalpha + IL-1beta	0.4
Primary Th1 rest	3.5	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.3	Small airway epithelium none	0.0
Primary Tr1 rest	2.5	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	9.3	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	12.4	Coronary artery SMC TNFalpha + IL-1beta	0.0

CD8 lymphocyte act	1.2	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	4.7	Astrocytes TNFalpha + IL-1beta	0.5
Secondary CD8 lymphocyte act	0.7	KU-812 (Basophil) rest	1.4
CD4 lymphocyte none	0.5	KU-812 (Basophil) PMA/ionomycin	4.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	1.2	CCD1106 (Keratinocytes) none	0.4
LAK cells rest	2.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	7.7
LAK cells IL-2	7.5	Liver cirrhosis	0.1
LAK cells IL-2+IL-12	7.1	NCI-H292 none	0.1
LAK cells IL-2+IFN gamma	9.5	NCI-H292 IL-4	0.3
LAK cells IL-2+ IL-18	11.7	NCI-H292 IL-9	0.4
LAK cells PMA/ionomycin	6.9	NCI-H292 IL-13	0.5
NK Cells IL-2 rest	11.3	NCI-H292 IFN gamma	2.0
Two Way MLR 3 day	14.0	HPAEC none	0.0
Two Way MLR 5 day	5.3	HPAEC TNF alpha + IL-1 beta	2.9
Two Way MLR 7 day	1.7	Lung fibroblast none	0.0
PBMC rest	0.4	Lung fibroblast TNF alpha + IL-1 beta	9.5
PBMC PWM	2.7	Lung fibroblast IL-4	0.0
PBMC PHA-L	1.1	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	3.6
B lymphocytes PWM	0.9	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.5	Dermal fibroblast CCD1070 TNF alpha	1.1
EOL-1 dbcAMP	0.6	Dermal fibroblast CCD1070 IL-1 beta	0.8
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	6.3
Dendritic cells none	0.6	Dermal fibroblast IL-4	0.3

Dendritic cells LPS	20.0	Dermal Fibroblasts rest	0.1
Dendritic cells anti-CD40	0.3	Neutrophils TNFa+LPS	0.1
Monocytes rest	2.7	Neutrophils rest	0.3
Monocytes LPS	100.0	Colon	0.5
Macrophages rest	0.9	Lung	0.9
Macrophages LPS	36.9	Thymus	1.0
HUVEC none	0.0	Kidney	0.4
HUVEC starved	0.1		

Table YI. Panel 4D

Tissue Name	Rel. Exp.(%) Ag1980, Run 161683379	Tissue Name	Rel. Exp.(%) Ag1980, Run 161683379
Secondary Th1 act	1.0	HUVEC IL-1beta	0.1
Secondary Th2 act	52.5	HUVEC IFN gamma	2.4
Secondary Tr1 act	9.4	HUVEC TNF alpha + IFN gamma	29.5
Secondary Th1 rest	8.9	HUVEC TNF alpha + IL4	1.0
Secondary Th2 rest	1.3	HUVEC IL-11	0.0
Secondary Tr1 rest	10.5	Lung Microvascular EC none	0.1
Primary Th1 act	4.5	Lung Microvascular EC TNFalpha + IL-1beta	1.2
Primary Th2 act	3.4	Microvascular Dermal EC none	0.3
Primary Tr1 act	4.5	Microvascular Dermal EC TNFalpha + IL-1beta	1.2
Primary Th1 rest	34.2	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	3.5	Small airway epithelium none	0.0
Primary Tr1 rest	5.8	Small airway epithelium TNFalpha + IL-1beta	0.2
CD45RA CD4 lymphocyte act	11.1	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	11.7	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	1.2	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	9.4	Astrocytes TNFalpha + IL-1beta	1.7
Secondary CD8 lymphocyte act	2.1	KU-812 (Basophil) rest	0.0

CD4 lymphocyte none	0.8	KU-812 (Basophil) PMA/ionomycin	15.1
2ry Th1/Th2/Tr1_anti-CD95 CH11	1.8	CCD1106 (Keratinocytes) none	0.8
LAK cells rest	6.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	2.6
LAK cells IL-2	19.5	Liver cirrhosis	0.3
LAK cells IL-2+IL-12	12.4	Lupus kidney	0.2
LAK cells IL-2+IFN gamma	53.2	NCI-H292 none	0.7
LAK cells IL-2+ IL-18	61.1	NCI-H292 IL-4	1.6
LAK cells PMA/ionomycin	12.4	NCI-H292 IL-9	2.2
NK Cells IL-2 rest	14.7	NCI-H292 IL-13	0.4
Two Way MLR 3 day	29.1	NCI-H292 IFN gamma	6.2
Two Way MLR 5 day	7.9	HPAEC none	0.0
Two Way MLR 7 day	3.4	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	1.3	Lung fibroblast none	0.0
PBMC PWM	17.8	Lung fibroblast TNF alpha + IL-1 beta	17.9
PBMC PHA-L	3.3	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	0.0
B lymphocytes PWM	12.6	Lung fibroblast IFN gamma	11.3
B lymphocytes CD40L and IL-4	2.7	Dermal fibroblast CCD1070 rest	0.1
EOL-1 dbcAMP	0.1	Dermal fibroblast CCD1070 TNF alpha	2.2
EOL-1 dbcAMP PMA/ionomycin	0.1	Dermal fibroblast CCD1070 IL-1 beta	0.5
Dendritic cells none	1.1	Dermal fibroblast IFN gamma	9.1
Dendritic cells LPS	48.6	Dermal fibroblast IL-4	0.4
Dendritic cells anti-CD40	0.4	IBD Colitis 2	0.1
Monocytes rest	5.7	IBD Crohn's	0.3
Monocytes LPS	100.0	Colon	3.6

Macrophages rest	1.3	Lung	3.6
Macrophages LPS	79.6	Thymus	0.8
HUVEC none	0.0	Kidney	3.0
HUVEC starved	0.2		

Table YJ. Panel 5 Islet

Tissue Name	Rel. Exp.(%) Ag3909, Run 242413199	Tissue Name	Rel. Exp.(%) Ag3909, Run 242413199
97457_Patient-02go_adipose	90.1	94709_Donor 2 AM - A_adipose	0.0
97476_Patient-07sk_skeletal muscle	22.2	94710_Donor 2 AM - B_adipose	1.8
97477_Patient-07ut_uterus	18.2	94711_Donor 2 AM - C_adipose	0.0
97478_Patient-07pl_placenta	53.6	94712_Donor 2 AD - A_adipose	1.6
99167_Bayer Patient 1	100.0	94713_Donor 2 AD - B_adipose	1.2
97482_Patient-08ut_uterus.	44.8	94714_Donor 2 AD - C_adipose	0.0
97483_Patient-08pl_placenta	74.2	94742_Donor 3 U - A_Mesenchymal Stem Cells	3.1
97486_Patient-09sk_skeletal muscle	9.3	94743_Donor 3 U - B_Mesenchymal Stem Cells	3.7
97487_Patient-09ut_uterus	24.1	94730_Donor 3 AM - A_adipose	1.7
97488_Patient-09pl_placenta	44.1	94731_Donor 3 AM - B_adipose	3.6
97492_Patient-10ut_uterus	47.0	94732_Donor 3 AM - C_adipose	0.0
97493_Patient-10pl_placenta	68.3	94733_Donor 3 AD - A_adipose	0.0
97495_Patient-11go_adipose	17.4	94734_Donor 3 AD - B_adipose	0.0
97496_Patient-11sk_skeletal muscle	22.8	94735_Donor 3 AD - C_adipose	0.0
97497_Patient-11ut_uterus	24.3	77138_Liver_HepG2untreated	1.8
97498_Patient-11pl_placenta	15.4	73556_Heart_Cardiac stromal cells (primary)	0.0
97500_Patient-12go_adipose	62.4	81735_Small Intestine	97.3
97501_Patient-	79.6	72409_Kidney_Proximal	5.6

12sk_skeletal muscle		Convulated Tubule	
97502_Patient-12ut_uterus	30.6	82685_Small intestine_Duodenum	54.3
97503_Patient-12pl_placenta	29.7	90650_Adrenal_Adrenocortical adenoma	10.0
94721_Donor 2 U - A_Mesenchymal Stem Cells	0.0	72410_Kidney_HRCE	8.3
94722_Donor 2 U - B_Mesenchymal Stem Cells	0.0	72411_Kidney_HRE	28.3
94723_Donor 2 U - C_Mesenchymal Stem Cells	0.0	73139_Uterus_Uterine smooth muscle cells	0.0

AI_comprehensive panel_v1.0 Summary: Ag1980 Two experiments with the same probe and primer produce results that are in excellent agreement, with highest expression in normal tissue adjacent to psoriasis (CTs=30.5-31.2). This target is induced in bone tissue, synovial fluid, synovial fluid cells and synovium from arthritis patients (rheumatoid-RA and osteoarthritis-OA); In addition, the expression of this transcript in these samples from normal patients is much lower. Other tissues including skin and lung also express this transcript. However, a consistent expression in diseased tissue, as compared to adjacent tissue or normal lung, is not apparent. This may be due to contamination with activated monocytes which highly express this transcript (see panel 4.1D)

CNS_neurodegeneration_v1.0 Summary: Ag3909 This panel does not show differential expression of the CG94235-01 gene in Alzheimer's disease. However, this expression profile confirms the presence of this gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system.

General_screening_panel_v1.4 Summary: Ag3909 Two experiments with the same probe and primer produce results that are in excellent agreement, with highest expression of the CG94235-01 gene in a gastric cancer cell line (CTs=23.6-24.4). Thus, expression of this gene could be used as to differentiate this sample from other samples on this panel and as a marker of gastric cancer. This gene encodes a putative thymidylate kinase, a DNA synthesis enzyme necessary for cell growth. Thus, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of gastric cancer.

Among tissues with metabolic function, this gene is expressed at moderate to low levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic disorders.

- 5 Dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

In addition, this gene is expressed at much higher levels in fetal lung, liver and skeletal muscle tissue (CTs=28-30) when compared to expression in the adult counterpart (CTs=32.5-35). Thus, expression of this gene may be used to differentiate between the fetal and adult source of these tissues.

This gene is also expressed at moderate to low levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 1.3D Summary: Ag1980 Highest expression of the CG94235-01 gene in this panel is seen in a gastric cancer cell line (CT=26). Overall, expression is in reasonable agreement with the results in Panel 1.4. Moderate to low levels of expression are seen in metabolic tissues including adipose, adult and fetal liver, skeletal muscle, heart, pituitary, thyroid, adrenal and pituitary. Moderate to low levels of expression are seen in all CNS regions examined.

In addition, higher levels of expression are seen in fetal liver (CT=30.2) when compared to expression in adult liver (CT=33.7). Thus, expression of this gene could be used to differentiate between the adult and fetal sources of this tissue.

Panel 2D Summary: Ag1980 Highest expression of the CG94235-01 gene is seen in normal bladder (CT=27.3). In addition, higher levels of expression are seen in ovarian, bladder and lung cancers when compared to expression in normal adjacent tissue. Thus, expression of this gene could be used as a marker of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of ovarian, bladder and lung cancers.

Panel 4.1D Summary: Ag3909 Highest expression of the CG94235-01 gene is seen in LPS treated monocytes. (CT=25.4). Prominent levels of expression are also seen in

LPS activated macrophages and dendritic cells. This transcript may encode a protein that is important in the normal regulation of cytokines. Inappropriate regulation of the protein encoded by this gene may result in the enhanced and uncontrolled expression of inflammatory cytokines. Therefore, designing therapies that could regulate the expression (such as antisense therapies) or function of the protein encoded by this gene may be important in the treatment of osteoarthritis and rheumatoid arthritis as well as other diseases.

Panel 4D Summary: Ag1980 The expression profile of the CG94235-01 gene in this panel is similar to the expression pattern seen in Panel 4.1D using the Ag3909 probe and primer set, except it is also expressed in LAK cells. This may reflect differences between panel 4 and 4.1 or slightly different amplicons. Please see panel 4.1D for discussion of utility of this gene in inflammation.

Panel 5 Islet Summary: Ag3909 Highest expression of the CG94235-01 gene is seen in islet cells (CT=33.4). Low but significant levels of expression are seen in other metabolic tissues, including adipose, placenta and skeletal muscle. Please see Panel 1.4 for discussion of utility of this gene in metabolic disease.

General oncology screening panel_v_2.4 Summary: Ag3909 Results from one experiment with the CG94235-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Z. CG94235-02: Splice Variant of CG94235-01

Expression of gene CG94235-02 was assessed using the primer-probe sets Ag3909 and Ag6052, described in Tables ZA and ZB. Results of the RTQ-PCR runs are shown in Tables ZC, ZD, ZE, ZF, ZG and ZH.

Table ZA. Probe Name Ag3909

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-caggtgccacgtctaactagat-3'	22	1529	299
Probe	TET-5'-tggtgtttgaaacatctacatccacca-3'-TAMRA	27	1498	300
Reverse	5'-gaaatttgggaacactgcataa-3'	22	1472	301

Table ZB. Probe Name Ag6052

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-cctaggcgccggttttg-3'	16	2242	302

Probe	TET-5'-agtgtacctcctttattctgaagcccg-3'-TAMRA	28	2210	303
Reverse	5'-gtcgaccaggtcaagcact-3'	19	2188	304

Table ZC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3909, Run 212248424	Tissue Name	Rel. Exp.(%) Ag3909, Run 212248424
AD 1 Hippo	11.7	Control (Path) 3 Temporal Ctx	5.6
AD 2 Hippo	41.5	Control (Path) 4 Temporal Ctx	33.7
AD 3 Hippo	7.6	AD 1 Occipital Ctx	12.5
AD 4 Hippo	7.2	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	75.3	AD 3 Occipital Ctx	1.1
AD 6 Hippo	36.1	AD 4 Occipital Ctx	23.7
Control 2 Hippo	19.1	AD 5 Occipital Ctx	29.9
Control 4 Hippo	12.9	AD 6 Occipital Ctx	35.4
Control (Path) 3 Hippo	12.2	Control 1 Occipital Ctx	9.1
AD 1 Temporal Ctx	9.5	Control 2 Occipital Ctx	69.7
AD 2 Temporal Ctx	36.3	Control 3 Occipital Ctx	18.9
AD 3 Temporal Ctx	3.7	Control 4 Occipital Ctx	3.9
AD 4 Temporal Ctx	31.9	Control (Path) 1 Occipital Ctx	71.2
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	31.9
AD 5 SupTemporal Ctx	56.6	Control (Path) 3 Occipital Ctx	4.3
AD 6 Inf Temporal Ctx	22.4	Control (Path) 4 Occipital Ctx	23.0
AD 6 Sup Temporal Ctx	32.3	Control 1 Parietal Ctx	8.5
Control 1 Temporal Ctx	5.3	Control 2 Parietal Ctx	38.2
Control 2 Temporal Ctx	42.9	Control 3 Parietal Ctx	29.1
Control 3 Temporal Ctx	13.0	Control (Path) 1 Parietal Ctx	73.2
Control 4 Temporal Ctx	4.4	Control (Path) 2 Parietal Ctx	44.4
Control (Path) 1 Temporal Ctx	58.2	Control (Path) 3 Parietal Ctx	7.3
Control (Path) 2 Temporal Ctx	41.8	Control (Path) 4 Parietal Ctx	49.0

Table ZD. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag3909, Run 217235826	Rel. Exp.(%) Ag3909, Run 219173644	Tissue Name	Rel. Exp.(%) Ag3909, Run 217235826	Rel. Exp.(%) Ag3909, Run 219173644
Adipose	0.6	0.6	Renal ca. TK-10	0.0	0.0
Melanoma* Hs688(A).T	0.0	0.0	Bladder	9.7	11.7
Melanoma* Hs688(B).T	0.0	0.0	Gastric ca. (liver met.) NCI-N87	100.0	100.0
Melanoma* M14	0.3	0.4	Gastric ca. KATO III	0.5	0.6
Melanoma* LOXIMVI	0.1	0.2	Colon ca. SW-948	0.5	0.5
Melanoma* SK-MEL-5	0.2	0.2	Colon ca. SW480	0.1	0.1
Squamous cell carcinoma SCC-4	0.5	0.5	Colon ca.* (SW480 met) SW620	0.0	0.0
Testis Pool	0.6	0.5	Colon ca. HT29	0.1	0.1
Prostate ca.* (bone met) PC-3	0.1	0.1	Colon ca. HCT-116	0.3	0.3
Prostate Pool	0.3	0.3	Colon ca. CaCo-2	0.1	0.0
Placenta	0.4	0.5	Colon cancer tissue	0.6	0.5
Uterus Pool	0.2	0.2	Colon ca. SW1116	0.2	0.1
Ovarian ca. OVCAR-3	0.2	0.2	Colon ca. Colo-205	0.0	0.0
Ovarian ca. SK-OV-3	0.5	0.4	Colon ca. SW-48	0.0	0.0
Ovarian ca. OVCAR-4	0.1	0.1	Colon Pool	0.4	0.5
Ovarian ca. OVCAR-5	0.6	0.5	Small Intestine Pool	0.4	0.4
Ovarian ca. IGROV-1	0.0	0.0	Stomach Pool	0.1	0.3
Ovarian ca. OVCAR-8	0.0	0.0	Bone Marrow Pool	0.2	0.2
Ovary	0.8	0.6	Fetal Heart	0.2	0.2
Breast ca. MCF-7	0.0	0.0	Heart Pool	0.5	0.5
Breast ca. MDA-MB-	0.2	0.2	Lymph Node Pool	0.4	0.4

231					
Breast ca. BT 549	2.3	3.1	Fetal Skeletal Muscle	0.2	0.2
Breast ca. T47D	1.1	1.0	Skeletal Muscle Pool	1.7	1.9
Breast ca. MDA-N	0.2	0.2	Spleen Pool	2.3	2.6
Breast Pool	0.3	0.4	Thymus Pool	0.5	0.5
Trachea	0.9	1.0	CNS cancer (glio/astro) U87-MG	0.0	0.0
Lung	0.1	0.1	CNS cancer (glio/astro) U-118-MG	1.0	1.3
Fetal Lung	0.9	0.9	CNS cancer (neuro;met) SK-N-AS	1.2	1.5
Lung ca. NCI-N417	0.0	0.0	CNS cancer (astro) SF-539	0.3	0.4
Lung ca. LX-1	0.0	0.0	CNS cancer (astro) SNB-75	0.1	0.1
Lung ca. NCI-H146	0.9	1.1	CNS cancer (glio) SNB-19	0.0	0.0
Lung ca. SHP-77	0.5	0.5	CNS cancer (glio) SF-295	0.2	0.2
Lung ca. A549	0.0	0.0	Brain (Amygdala) Pool	0.6	0.5
Lung ca. NCI-H526	0.0	0.0	Brain (cerebellum)	0.1	0.2
Lung ca. NCI-H23	0.0	0.0	Brain (fetal)	0.8	0.7
Lung ca. NCI-H460	0.0	0.0	Brain (Hippocampus) Pool	0.5	0.5
Lung ca. HOP-62	0.0	0.0	Cerebral Cortex Pool	0.7	0.7
Lung ca. NCI-H522	0.0	0.0	Brain (Substantia nigra) Pool	0.8	0.7
Liver	0.1	0.1	Brain (Thalamus) Pool	0.8	0.9
Fetal Liver	4.7	4.1	Brain (whole)	1.3	1.3
Liver ca. HepG2	0.0	0.0	Spinal Cord Pool	0.4	0.5
Kidney Pool	0.7	0.8	Adrenal Gland	0.7	0.8
Fetal Kidney	0.2	0.2	Pituitary gland Pool	0.3	0.3

Renal ca. 786-0	0.0	0.0	Salivary Gland	0.5	0.4
Renal ca. A498	0.2	0.2	Thyroid (female)	0.4	0.2
Renal ca. ACHN	0.0	0.0	Pancreatic ca. CAPAN2	0.3	0.3
Renal ca. UO-31	0.1	0.0	Pancreas Pool	0.4	0.3

Table ZE. General_screening_panel_v1.5

Tissue Name	Rel. Exp.(%) Ag6052, Run 228746663	Tissue Name	Rel. Exp.(%) Ag6052, Run 228746663
Adipose	1.2	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	12.9
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	100.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.4
Squamous cell carcinoma SCC-4	2.7	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.4
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.4
Uterus Pool	0.0	Colon ca. SW1116	0.9
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.6
Ovarian ca. OVCAR-5	0.6	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.4

Ovary	0.7	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.3
Breast ca. MDA-MB-231	0.5	Lymph Node Pool	0.0
Breast ca. BT 549	1.5	Fetal Skeletal Muscle	1.5
Breast ca. T47D	0.0	Skeletal Muscle Pool	2.1
Breast ca. MDA-N	0.0	Spleen Pool	3.7
Breast Pool	0.9	Thymus Pool	1.2
Trachea	1.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.8
Fetal Lung	0.4	CNS cancer (neuro:met) SK-N-AS	2.3
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.4	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	1.9
Lung ca. A549	0.0	Brain (Amygdala) Pool	1.3
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	1.2
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	1.8
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.4
Fetal Liver	2.8	Brain (whole)	0.0
Liver ca. HepG2	0.0	Spinal Cord Pool	1.4
Kidney Pool	0.0	Adrenal Gland	0.4
Fetal Kidney	1.6	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.6	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.3
Renal ca. UO-31	0.0	Pancreas Pool	0.6

Table ZF. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3909, Run 170127176	Rel. Exp.(%) Ag6052, Run 226202125	Tissue Name	Rel. Exp.(%) Ag3909, Run 170127176	Rel. Exp.(%) Ag6052, Run 226202125
Secondary Th1 act	0.9	0.0	HUVEC IL-1beta	0.1	0.0

Secondary Th2 act	32.3	27.7	HUVEC IFN gamma	0.9	0.5
Secondary Tr1 act	3.6	0.8	HUVEC TNF alpha + IFN gamma	10.5	18.3
Secondary Th1 rest	5.5	5.9	HUVEC TNF alpha + IL4	0.4	0.6
Secondary Th2 rest	2.0	1.0	HUVEC IL-11	0.0	2.7
Secondary Tr1 rest	4.8	4.3	Lung Microvascular EC none	0.2	0.0
Primary Th1 act	1.3	1.1	Lung Microvascular EC TNFalpha + IL-1beta	0.6	0.0
Primary Th2 act	2.0	2.2	Microvascular Dermal EC none	0.1	0.0
Primary Tr1 act	1.1	0.4	Microvascular Dermal EC TNFalpha + IL-1beta	0.4	0.0
Primary Th1 rest	3.5	4.4	Bronchial epithelium TNFalpha + IL1beta	0.0	0.0
Primary Th2 rest	0.3	2.1	Small airway epithelium none	0.0	0.0
Primary Tr1 rest	2.5	0.3	Small airway epithelium TNFalpha + IL-1beta	0.0	0.0
CD45RA CD4 lymphocyte act	9.3	4.2	Coronary artery SMC rest	0.0	0.0
CD45RO CD4 lymphocyte act	12.4	11.7	Coronary artery SMC TNFalpha + IL-1beta	0.0	0.0
CD8 lymphocyte act	1.2	0.6	Astrocytes rest	0.0	0.0
Secondary CD8 lymphocyte rest	4.7	12.9	Astrocytes TNFalpha + IL-1beta	0.5	0.3
Secondary CD8 lymphocyte act	0.7	0.4	KU-812 (Basophil) rest	1.4	1.1
CD4 lymphocyte none	0.5	0.0	KU-812 (Basophil) PMA/ionomycin	4.7	6.3
2ry Th1/Th2/Tr1_anti-CD95 CH11	1.2	1.9	CCD1106 (Keratinocytes) none	0.4	0.7
LAK cells rest	2.8	2.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	7.7	6.5
LAK cells IL-2	7.5	8.4	Liver cirrhosis	0.1	0.0
LAK cells IL-	7.1	3.2	NCI-H292 none	0.1	0.0

2+IL-12					
LAK cells IL-2+IFN gamma	9.5	2.7	NCI-H292 IL-4	0.3	0.0
LAK cells IL-2+IL-18	11.7	7.1	NCI-H292 IL-9	0.4	0.0
LAK cells PMA/ionomycin	6.9	5.5	NCI-H292 IL-13	0.5	0.0
NK Cells IL-2 rest	11.3	13.4	NCI-H292 IFN gamma	2.0	0.4
Two Way MLR 3 day	14.0	12.7	HPAEC none	0.0	0.6
Two Way MLR 5 day	5.3	3.6	HPAEC TNF alpha + IL-1 beta	2.9	0.4
Two Way MLR 7 day	1.7	5.2	Lung fibroblast none	0.0	0.6
PBMC rest	0.4	0.0	Lung fibroblast TNF alpha + IL-1 beta	9.5	18.6
PBMC PWM	2.7	1.5	Lung fibroblast IL-4	0.0	5.2
PBMC PHA-L	1.1	4.2	Lung fibroblast IL-9	0.0	0.0
Ramos (B cell) none	0.0	0.0	Lung fibroblast IL-13	0.0	0.0
Ramos (B cell) ionomycin	0.0	0.0	Lung fibroblast IFN gamma	3.6	5.7
B lymphocytes PWM	0.9	0.7	Dermal fibroblast CCD1070 rest	0.0	0.0
B lymphocytes CD40L and IL-4	0.5	1.5	Dermal fibroblast CCD1070 TNF alpha	1.1	1.6
EOL-1 dbcAMP	0.6	0.3	Dermal fibroblast CCD1070 IL-1 beta	0.8	1.6
EOL-1 dbcAMP PMA/ionomycin	0.0	0.0	Dermal fibroblast IFN gamma	6.3	2.9
Dendritic cells none	0.6	0.0	Dermal fibroblast IL-4	0.3	0.0
Dendritic cells LPS	20.0	15.1	Dermal Fibroblasts rest	0.1	0.0
Dendritic cells anti-CD40	0.3	0.4	Neutrophils TNFa+LPS	0.1	1.6
Monocytes rest	2.7	2.8	Neutrophils rest	0.3	0.0
Monocytes LPS	100.0	100.0	Colon	0.5	3.0
Macrophages rest	0.9	0.8	Lung	0.9	1.4
Macrophages LPS	36.9	24.3	Thymus	1.0	1.3
HUVEC none	0.0	0.0	Kidney	0.4	0.0
HUVEC starved	0.1	0.0			

Table ZG. Panel 5 Islet

Tissue Name	Rel. Exp.(%) Ag3909, Run 242413199	Tissue Name	Rel. Exp.(%) Ag3909, Run 242413199
97457_Patient-02go_adipose	90.1	94709_Donor 2 AM - A_adipose	0.0
97476_Patient-07sk_skeletal muscle	22.2	94710_Donor 2 AM - B_adipose	1.8
97477_Patient-07ut_uterus	18.2	94711_Donor 2 AM - C_adipose	0.0
97478_Patient-07pl_placenta	53.6	94712_Donor 2 AD - A_adipose	1.6
99167_Bayer Patient 1	100.0	94713_Donor 2 AD - B_adipose	1.2
97482_Patient-08ut_uterus	44.8	94714_Donor 2 AD - C_adipose	0.0
97483_Patient-08pl_placenta	74.2	94742_Donor 3 U - A_Mesenchymal Stem Cells	3.1
97486_Patient-09sk_skeletal muscle	9.3	94743_Donor 3 U - B_Mesenchymal Stem Cells	3.7
97487_Patient-09ut_uterus	24.1	94730_Donor 3 AM - A_adipose	1.7
97488_Patient-09pl_placenta	44.1	94731_Donor 3 AM - B_adipose	3.6
97492_Patient-10ut_uterus	47.0	94732_Donor 3 AM - C_adipose	0.0
97493_Patient-10pl_placenta	68.3	94733_Donor 3 AD - A_adipose	0.0
97495_Patient-11go_adipose	17.4	94734_Donor 3 AD - B_adipose	0.0
97496_Patient-11sk_skeletal muscle	22.8	94735_Donor 3 AD - C_adipose	0.0
97497_Patient-11ut_uterus	24.3	77138_Liver_HepG2untreated	1.8
97498_Patient-11pl_placenta	15.4	73556_Heart_Cardiac stromal cells (primary)	0.0
97500_Patient-12go_adipose	62.4	81735_Small Intestine	97.3
97501_Patient-12sk_skeletal muscle	79.6	72409_Kidney_Proximal Convolutd Tubule	5.6
97502_Patient-12ut_uterus	30.6	82685_Small intestine_Duodenum	54.3
97503_Patient-12pl_placenta	29.7	90650_Adrenal_Adrenocortical adenoma	10.0
94721_Donor 2 U - A_Mesenchymal Stem Cells	0.0	72410_Kidney_HRCE	8.3
94722_Donor 2 U - B_Mesenchymal Stem Cells	0.0	72411_Kidney_HRE	28.3
94723_Donor 2 U - C_Mesenchymal Stem Cells	0.0	73139_Uterus_Uterine smooth muscle cells	0.0

Table ZH. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3909, Run	Tissue Name	Rel. Exp.(%)
-------------	-----------------------------	-------------	--------------

	268143664		Ag3909, Run 268143664
Colon cancer 1	0.0	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	0.0
Colon cancer 2	0.0	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	0.0
Colon cancer 3	0.1	Adenocarcinoma of the prostate 2	100.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	0.0
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	0.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	0.0	Kidney NAT 1	0.0
Melanoma 2	0.0	Kidney cancer 2	0.0
Melanoma 3	0.0	Kidney NAT 2	0.0
metastatic melanoma 4	0.0	Kidney cancer 3	0.0
metastatic melanoma 5	0.0	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag3909 This panel does not show differential expression of the CG94235-01 gene in Alzheimer's disease. However, this expression profile confirms the presence of this gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system. A second experiment with the probe and primer set Ag6052, which is specific for the CG94235-02 variant, shows low/undetectable levels of expression in all the samples on this panel (CTs>35). (Data not shown.)

General_screening_panel_v1.4 Summary: Ag3909 Two experiments with the same probe and primer produce results that are in excellent agreement, with highest expression of the CG94235-01 gene in a gastric cancer cell line (CTs=23.6-24.4). Thus, expression of this gene could be used as to differentiate this sample from other samples on

this panel and as a marker of gastric cancer. This gene encodes a putative thymidylate kinase, a DNA synthesis enzyme necessary for cell growth. Thus, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of gastric cancer.

5 Among tissues with metabolic function, this gene is expressed at moderate to low levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that
10 disregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

 In addition, this gene is expressed at much higher levels in fetal lung, liver and skeletal muscle tissue (CTs=28-30) when compared to expression in the adult counterpart (CTs=32.5-35). Thus, expression of this gene may be used to differentiate between the fetal and adult source of these tissues.

15 This gene is also expressed at moderate to low levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

20 **General_screening_panel_v1.5 Summary:** Ag6052 Expression of the CG94235-02 variant appears to be restricted to a gastric cancer cell line (CT=31.2) and normal bladder (CT=34.2) in this panel. Thus, expression of this gene could be used to differentiate between these samples and other samples on this panel and as a marker to detect the presence of gastric cancer. Furthermore, therapeutic modulation of the
25 expression or function of this gene may be effective in the treatment of gastric cancer.

Panel 4.1D Summary: Ag3909/Ag6502 Two experiments with two different probe and primer sets produce results that are in very good agreement. Highest expression of the CG94235-01 gene is seen in LPS treated monocytes. (Ag3909 CT=25.4; Ag6502 CT=31.1). Prominent levels of expression are also seen in LPS activated macrophages and
30 dendritic cells. This transcript may encode a protein that is important in the normal regulation of cytokines. Inappropriate regulation of the protein encoded by this gene may result in the enhanced and uncontrolled expression of inflammatory cytokines. Therefore,

designing therapies that could regulate the expression (such as antisense therapies) or function of the protein encoded by CG94325-01 this gene may be important in the treatment of osteoarthritis and rheumatoid arthritis as well as other diseases.

Panel 5 Islet Summary: Ag3909 Highest expression of the CG94235-01 gene is seen in islet cells (CT=33.4). Low but significant levels of expression are seen in other metabolic tissues, including adipose, placenta and skeletal muscle. Please see Panel 1.4 for discussion of utility of this gene in metabolic disease.

General oncology screening panel_v_2.4 Summary: Ag3909 Results from one experiment with the CG94235-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run. A second experiment with the probe and primer set Ag6052, which is specific for the CG94235-02 variant, shows low/undetectable levels of expression in all the samples on this panel (CTs>35). (Data not shown.)

AA. CG94692-01 and CG94692-02: Carnitine/Acylcarnitin Translocase

Expression of gene CG94692-01 and full length physical clone CG94692-02 was assessed using the primer-probe set Ag3941, described in Table AAA. Results of the RTQ-PCR runs are shown in Tables AAB, AAC, AAD and AAE. Please note that CG94692-02 represents a full-length physical clone of the CG94692-01 gene, validating the prediction of the gene sequence.

Table AAA. Probe Name Ag3941

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggcttcttcaagggaatgag-3'	20	160	305
Probe	TET-5'-ttgccagcatagctgtggtcaactct-3'-TAMRA	26	188	306
Reverse	5'-tgttgctatagaccccaaacag-3'	22	217	307

Table AAB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3941, Run 212344932	Tissue Name	Rel. Exp.(%) Ag3941, Run 212344932
AD 1 Hippo	6.3	Control (Path) 3 Temporal Ctx	4.2
AD 2 Hippo	16.5	Control (Path) 4 Temporal Ctx	28.9
AD 3 Hippo	3.0	AD 1 Occipital Ctx	2.0
AD 4 Hippo	6.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	100.0	AD 3 Occipital Ctx	0.6
AD 6 Hippo	10.6	AD 4 Occipital Ctx	17.1

Control 2 Hippo	17.3	AD 5 Occipital Ctx	28.7
Control 4 Hippo	8.0	AD 6 Occipital Ctx	6.4
Control (Path) 3 Hippo	4.5	Control 1 Occipital Ctx	2.1
AD 1 Temporal Ctx	6.7	Control 2 Occipital Ctx	39.5
AD 2 Temporal Ctx	27.9	Control 3 Occipital Ctx	17.4
AD 3 Temporal Ctx	3.9	Control 4 Occipital Ctx	5.7
AD 4 Temporal Ctx	18.9	Control (Path) 1 Occipital Ctx	51.4
AD 5 Inf Temporal Ctx	39.8	Control (Path) 2 Occipital Ctx	11.5
AD 5 Sup Temporal Ctx	25.0	Control (Path) 3 Occipital Ctx	1.6
AD 6 Inf Temporal Ctx	13.8	Control (Path) 4 Occipital Ctx	16.7
AD 6 Sup Temporal Ctx	11.9	Control 1 Parietal Ctx	4.1
Control 1 Temporal Ctx	5.2	Control 2 Parietal Ctx	13.7
Control 2 Temporal Ctx	33.9	Control 3 Parietal Ctx	13.6
Control 3 Temporal Ctx	6.0	Control (Path) 1 Parietal Ctx	36.9
Control 3 Temporal Ctx	10.7	Control (Path) 2 Parietal Ctx	21.9
Control (Path) 1 Temporal Ctx	45.4	Control (Path) 3 Parietal Ctx	2.8
Control (Path) 2 Temporal Ctx	23.3	Control (Path) 4 Parietal Ctx	42.6

Table AAC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag3941, Run 219822166	Tissue Name	Rel. Exp.(%) Ag3941, Run 219822166
Adipose	6.3	Renal ca. TK-10	7.5
Melanoma* Hs688(A).T	25.3	Bladder	11.3
Melanoma* Hs688(B).T	8.1	Gastric ca. (liver met.) NCI-N87	13.8
Melanoma* M14	10.6	Gastric ca. KATO III	3.7
Melanoma* LOXIMVI	10.5	Colon ca. SW-948	8.2
Melanoma* SK-MEL-5	17.8	Colon ca. SW480	21.3
Squamous cell carcinoma SCC-4	1.3	Colon ca.* (SW480 met) SW620	14.0
Testis Pool	7.6	Colon ca. HT29	5.0

Prostate ca.* (bone met) PC-3	32.3	Colon ca. HCT-116	6.3
Prostate Pool	8.5	Colon ca. CaCo-2	12.0
Placenta	12.7	Colon cancer tissue	7.4
Uterus Pool	1.2	Colon ca. SW1116	19.5
Ovarian ca. OVCAR-3	16.4	Colon ca. Colo-205	1.1
Ovarian ca. SK-OV-3	7.2	Colon ca. SW-48	6.6
Ovarian ca. OVCAR-4	3.3	Colon Pool	12.9
Ovarian ca. OVCAR-5	55.1	Small Intestine Pool	8.2
Ovarian ca. IGROV-1	22.1	Stomach Pool	4.5
Ovarian ca. OVCAR-8	6.8	Bone Marrow Pool	4.6
Ovary	6.8	Fetal Heart	7.2
Breast ca. MCF-7	11.2	Heart Pool	5.6
Breast ca. MDA-MB-231	11.5	Lymph Node Pool	11.0
Breast ca. BT 549	9.1	Fetal Skeletal Muscle	15.5
Breast ca. T47D	100.0	Skeletal Muscle Pool	12.2
Breast ca. MDA-N	32.5	Spleen Pool	18.6
Breast Pool	9.1	Thymus Pool	23.0
Trachea	16.4	CNS cancer (glio/astro) U87-MG	19.8
Lung	1.0	CNS cancer (glio/astro) U-118-MG	15.6
Fetal Lung	29.7	CNS cancer (neuro;met) SK-N-AS	5.0
Lung ca. NCI-N417	2.9	CNS cancer (astro) SF-539	3.7
Lung ca. LX-1	18.3	CNS cancer (astro) SNB-75	13.7
Lung ca. NCI-H146	8.3	CNS cancer (glio) SNB-19	18.0
Lung ca. SHP-77	12.7	CNS cancer (glio) SF-295	14.6
Lung ca. A549	18.8	Brain (Amygdala) Pool	7.2
Lung ca. NCI-H526	5.2	Brain (cerebellum)	5.2
Lung ca. NCI-H23	24.1	Brain (fetal)	1.5
Lung ca. NCI-H460	6.3	Brain (Hippocampus) Pool	6.2
Lung ca. HOP-62	7.0	Cerebral Cortex Pool	8.4
Lung ca. NCI-H522	15.4	Brain (Substantia nigra) Pool	12.3
Liver	5.2	Brain (Thalamus) Pool	9.2
Fetal Liver	15.9	Brain (whole)	7.7
Liver ca. HepG2	9.5	Spinal Cord Pool	8.0
Kidney Pool	18.6	Adrenal Gland	9.5
Fetal Kidney	9.0	Pituitary gland Pool	4.3
Renal ca. 786-0	3.4	Salivary Gland	8.7
Renal ca. A498	1.5	Thyroid (female)	23.7

Renal ca. ACHN	5.5	Pancreatic ca. CAPAN2	6.2
Renal ca. UO-31	3.3	Pancreas Pool	17.0

Table AAD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3941, Run 170684829	Tissue Name	Rel. Exp.(%) Ag3941, Run 170684829
Secondary Th1 act	36.9	HUVEC IL-1beta	9.3
Secondary Th2 act	43.5	HUVEC IFN gamma	5.4
Secondary Tr1 act	47.3	HUVEC TNF alpha + IFN gamma	5.9
Secondary Th1 rest	46.0	HUVEC TNF alpha + IL4	8.7
Secondary Th2 rest	45.1	HUVEC IL-11	7.0
Secondary Tr1 rest	58.6	Lung Microvascular EC none	11.7
Primary Th1 act	31.6	Lung Microvascular EC TNFalpha + IL-1beta	14.2
Primary Th2 act	72.7	Microvascular Dermal EC none	6.0
Primary Tr1 act	47.0	Microvascular Dermal EC TNFalpha + IL-1beta	4.8
Primary Th1 rest	40.6	Bronchial epithelium TNFalpha + IL1beta	2.4
Primary Th2 rest	26.2	Small airway epithelium none	1.4
Primary Tr1 rest	100.0	Small airway epithelium TNFalpha + IL-1beta	4.1
CD45RA CD4 lymphocyte act	27.9	Coronary artery SMC rest	9.4
CD45RO CD4 lymphocyte act	45.1	Coronary artery SMC TNFalpha + IL-1beta	14.2
CD8 lymphocyte act	44.1	Astrocytes rest	4.2
Secondary CD8 lymphocyte rest	37.1	Astrocytes TNFalpha + IL-1beta	7.1
Secondary CD8 lymphocyte act	33.2	KU-812 (Basophil) rest	16.8
CD4 lymphocyte none	63.7	KU-812 (Basophil) PMA/ionomycin	11.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	85.9	CCD1106 (Keratinocytes) none	2.6
LAK cells rest	73.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.8
LAK cells IL-2	68.3	Liver cirrhosis	8.1
LAK cells IL-2+IL-12	42.9	NCI-H292 none	19.2
LAK cells IL-2+IFN gamma	35.8	NCI-H292 IL-4	17.9
LAK cells IL-2+ IL-18	59.5	NCI-H292 IL-9	38.2

LAK cells PMA/ionomycin	14.7	NCI-H292 IL-13	11.9
NK Cells IL-2 rest	81.2	NCI-H292 IFN gamma	26.1
Two Way MLR 3 day	63.3	HPAEC none	4.0
Two Way MLR 5 day	31.9	HPAEC TNF alpha + IL-1 beta	13.3
Two Way MLR 7 day	28.9	Lung fibroblast none	6.8
PBMC rest	26.8	Lung fibroblast TNF alpha + IL-1 beta	9.0
PBMC PWM	18.2	Lung fibroblast IL-4	7.6
PBMC PHA-L	33.0	Lung fibroblast IL-9	19.6
Ramos (B cell) none	7.2	Lung fibroblast IL-13	7.9
Ramos (B cell) ionomycin	20.6	Lung fibroblast IFN gamma	7.6
B lymphocytes PWM	25.0	Dermal fibroblast CCD1070 rest	11.3
B lymphocytes CD40L and IL-4	74.2	Dermal fibroblast CCD1070 TNF alpha	40.1
EOL-1 dbcAMP	2.5	Dermal fibroblast CCD1070 IL-1 beta	1.0
EOL-1 dbcAMP PMA/ionomycin	1.6	Dermal fibroblast IFN gamma	4.9
Dendritic cells none	23.8	Dermal fibroblast IL-4	15.4
Dendritic cells LPS	23.2	Dermal Fibroblasts rest	8.7
Dendritic cells anti-CD40	36.3	Neutrophils TNFa+LPS	2.9
Monocytes rest	31.6	Neutrophils rest	11.3
Monocytes LPS	16.4	Colon	40.9
Macrophages rest	30.1	Lung	37.6
Macrophages LPS	13.8	Thymus	69.7
HUVEC none	4.1	Kidney	85.9
HUVEC starved	5.8		

Table AAE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3941, Run 268035089	Tissue Name	Rel. Exp.(%) Ag3941, Run 268035089
Colon cancer 1	7.9	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	4.2	Bladder cancer NAT 3	0.0
Colon cancer 2	6.5	Bladder cancer NAT 4	2.0
Colon cancer NAT 2	7.9	Adenocarcinoma of the prostate 1	13.4
Colon cancer 3	8.0	Adenocarcinoma of the prostate 2	0.2
Colon cancer NAT 3	4.8	Adenocarcinoma of the prostate 3	7.4
Colon malignant cancer 4	9.3	Adenocarcinoma of the prostate 4	6.9
Colon normal adjacent tissue	2.7	Prostate cancer NAT 5	0.4

4			
Lung cancer 1	7.5	Adenocarcinoma of the prostate 6	2.9
Lung NAT 1	1.2	Adenocarcinoma of the prostate 7	3.4
Lung cancer 2	21.5	Adenocarcinoma of the prostate 8	1.3
Lung NAT 2	1.1	Adenocarcinoma of the prostate 9	15.8
Squamous cell carcinoma 3	15.7	Prostate cancer NAT 10	1.1
Lung NAT 3	0.4	Kidney cancer 1	16.6
metastatic melanoma 1	7.9	Kidney NAT 1	4.2
Melanoma 2	0.2	Kidney cancer 2	100.0
Melanoma 3	0.8	Kidney NAT 2	5.2
metastatic melanoma 4	19.2	Kidney cancer 3	16.0
metastatic melanoma 5	23.5	Kidney NAT 3	4.0
Bladder cancer 1	0.5	Kidney cancer 4	14.6
Bladder cancer NAT 1	0.0	Kidney NAT 4	18.3
Bladder cancer 2	1.9		

CNS_neurodegeneration_v1.0 Summary: Ag3941 The CG94692-01 gene

appears to be slightly downregulated in the brains of Alzheimer's patients, when compared to expression in normal control brains. Therefore, up-regulation of this gene or its protein product, or treatment with specific agonists for this receptor may be of use in reversing the dementia/memory loss and neuronal death associated with this disease.

General_screening_panel_v1.4 Summary: Ag3941 This gene is widely

expressed in this panel, with highest expression in a breast cancer cell line (CT=28.6). Prominent levels of expression are also seen in samples derived from prostate cancer, ovarian cancer and melanoma. Thus, expression of this gene could be used to differentiate these samples from the rest of the samples on this panel and as a marker of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of breast, ovarian, prostate and melanoma cancers. Overall, the widespread expression of this gene suggests a role for this gene in cell survival and growth.

Among tissues with metabolic function, this gene is expressed at moderate to low levels in pituitary, adipose, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

In addition, this gene is expressed at much higher levels in fetal lung (CT=30.5) when compared to expression in the adult counterpart (CT=35.2). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue.

This gene is also expressed at moderate to low levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 4.1D Summary: Ag3941 Highest expression of the CG94692-01 gene is seen in primary resting Tr1 cells (CT=30.9). The transcript is expressed at higher levels in lymphocytes, which is consistent with the expression in the thymus and lymph node on Panel 1.4. Therefore, therapeutics designed with this sequence or the protein it encodes could be important in regulating T cell activation and be important for immune modulation and in treating T and B cell mediated diseases such as asthma, allergy, COPD, arthritis, psoriasis, lupus and IBD.

General oncology screening panel_v_2.4 Summary: Ag3941 Highest expression of the CG94692-01 gene is seen in kidney cancer (CT=29.5). In addition, significant levels of expression are also seen in kidney cancers and lung cancers when compared to normal adjacent tissue. Thus, expression of this gene could be used to differentiate these samples from the rest of the samples on this panel and as a marker of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of kidney and lung cancers.

AB. CG94724-01: Acylcarnitine Translocase

Expression of gene CG94724-01 was assessed using the primer-probe set Ag4071, described in Table ABA. Results of the RTQ-PCR runs are shown in Table ABB.

Table ABA. Probe Name Ag4071

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctggtacattcaccacaggc-3'	20	363	308
Probe	TET-5'-cctggagaacccatcaagtccttggt-3'-TAMRA	26	392	309
Reverse	5'-acttggtttccctgaagaa-3'	20	430	310

Table ABB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4071, Run 171807825	Tissue Name	Rel. Exp.(%) Ag4071, Run 171807825
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	3.4	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	4.9	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	10.7	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0

NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	4.2	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	7.9
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	5.1
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	0.0
Macrophages rest	0.0	Lung	9.5
Macrophages LPS	0.0	Thymus	6.9
HUVEC none	0.0	Kidney	100.0
HUVEC starved	3.7		

CNS_neurodegeneration_v1.0 Summary: Ag4071 Expression of the CG94724-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

5 **General_screening_panel_v1.4 Summary:** Ag4071 Expression of the CG94724-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4071 Expression of the CG94724-01 gene is exclusively seen in kidney (CT=32.9). Therefore, expression of this gene can be used to distinguish kidney sample from other samples in this panel. In addition, therapeutic modulation of this gene through the use of small molecule target may be beneficial in the

treatment of autoimmune and inflammatory diseases that affect kidney, including lupus and glomerulonephritis.

The CG94724-01 gene codes for carnitine/acylcarnitine translocase (CACT) homologue. Carnitine-acylcarnitine translocase is 1 of 10 closely related mitochondrial-membrane carrier proteins that shuttle substrates between cytosol and the intramitochondrial matrix space. Deficiency in CACT causes defect in the co-transport of free and esterified carnitine across the inner mitochondrial membrane. Recently, Choong et al (2001, *Pediatr Dev Pathol* 4(6):573-9, PMID: 11826365) reported a case of lethal cardiac tachyarrhythmia in a newborn who died at 72 h of age from severe, intractable cardiac tachyarrhythmia, despite an improvement in his neurological and biochemical status caused due to CACT deficiency. Postmortem examination showed marked steatosis of myocardium, liver, and kidney. Thus, the CG94724-01 gene may also play a role in the pathology of disorders associated with CACT deficiency and therapeutic modulation of this gene product could be useful in the treatment of these disorders including lethal cardiac tachyarrhythmia.

Panel 5 Islet Summary: Ag4071 Expression of the CG94724-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4071 Expression of the CG94724-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

AC. CG94946-01 and CG94946-02 and CG94946-04 and CG94946-05 and CG94946-06 and CG94946-07: Agrin Precursor

Expression of gene CG94946-01 and variants CG94946-02, CG94946-04, CG94946-05, CG94946-06 and CG94946-07 was assessed using the primer-probe sets Ag3605 and Ag3974, described in Tables ACA and ACB. Results of the RTQ-PCR runs are shown in Tables ACC, ACD, ACE, ACF, ACG and ACH. Please note that variants CG94946-02, CG94946-04 and CG94946-07 correspond to the probe and primer set Ag3974 only.

Table ACA. Probe Name Ag3605

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gaccccaagtcagaactgttc-3'	21	3514	311

Probe	TET-5'-attgagagcaccctggacgacctctt-3'-TAMRA	26	3553	312
Reverse	5'-gaaatccttcttgacgtctgaa-3'	22	3585	313

Table ACB. Probe Name Ag3974

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gacaccaggatcttctttgtga-3'	22	379	314
Probe	TET-5'-catacctgtggccagccacaag-3'-TAMRA	23	413	315
Reverse	5'-gagttgagcatcagctcgtt-3'	20	436	316

Table ACC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3605, Run 210997601	Rel. Exp.(%) Ag3974, Run 212348647	Tissue Name	Rel. Exp.(%) Ag3605, Run 210997601	Rel. Exp.(%) Ag3974, Run 212348647
AD 1 Hippo	12.2	28.7	Control (Path) 3 Temporal Ctx	10.9	15.4
AD 2 Hippo	27.2	36.3	Control (Path) 4 Temporal Ctx	57.0	46.0
AD 3 Hippo	15.3	19.2	AD 1 Occipital Ctx	26.1	28.9
AD 4 Hippo	34.4	21.8	AD 2 Occipital Ctx (Missing)	0.0	0.0
AD 5 Hippo	100.0	80.7	AD 3 Occipital Ctx	12.0	19.5
AD 6 Hippo	32.3	42.6	AD 4 Occipital Ctx	24.7	21.0
Control 2 Hippo	37.9	42.3	AD 5 Occipital Ctx	44.4	47.6
Control 4 Hippo	20.7	34.9	AD 6 Occipital Ctx	13.5	14.7
Control (Path) 3 Hippo	7.5	12.4	Control 1 Occipital Ctx	10.0	20.4
AD 1 Temporal Ctx	28.7	32.1	Control 2 Occipital Ctx	51.8	55.1
AD 2 Temporal Ctx	31.9	31.2	Control 3 Occipital Ctx	17.3	22.5
AD 3 Temporal Ctx	17.4	20.2	Control 4 Occipital Ctx	14.9	22.2
AD 4 Temporal Ctx	31.2	24.0	Control (Path) 1 Occipital Ctx	90.8	71.2
AD 5 Inf Temporal Ctx	87.1	100.0	Control (Path) 2 Occipital Ctx	19.8	17.4
AD 5 Sup Temporal Ctx	51.4	58.6	Control (Path) 3 Occipital Ctx	7.4	15.4
AD 6 Inf Temporal Ctx	42.9	40.6	Control (Path) 4 Occipital Ctx	44.8	38.7
AD 6 Sup	51.1	35.1	Control 1 Parietal Ctx	14.6	18.2

Temporal Ctx					
Control 1 Temporal Ctx	17.3	19.8	Control 2 Parietal Ctx	53.6	67.4
Control 2 Temporal Ctx	50.3	48.3	Control 3 Parietal Ctx	18.6	21.2
Control 3 Temporal Ctx	23.7	17.8	Control (Path) 1 Parietal Ctx	76.3	51.4
Control 3 Temporal Ctx	20.3	25.0	Control (Path) 2 Parietal Ctx	36.6	32.3
Control (Path) 1 Temporal Ctx	78.5	63.7	Control (Path) 3 Parietal Ctx	12.0	11.9
Control (Path) 2 Temporal Ctx	50.3	43.5	Control (Path) 4 Parietal Ctx	64.2	58.6

Table ACD. General screening panel_v1.4

Tissue Name	Rel. Exp.(%) Ag3605, Run 213406184	Rel. Exp.(%) Ag3974, Run 217508632	Tissue Name	Rel. Exp.(%) Ag3605, Run 213406184	Rel. Exp.(%) Ag3974, Run 217508632
Adipose	1.4	1.5	Renal ca. TK-10	19.3	16.4
Melanoma* Hs688(A).T	2.6	3.2	Bladder	8.4	9.0
Melanoma* Hs688(B).T	4.6	4.2	Gastric ca. (liver met.) NCI-N87	87.7	80.7
Melanoma* M14	6.7	6.4	Gastric ca. KATO III	17.8	17.7
Melanoma* LOXIMVI	4.8	4.0	Colon ca. SW-948	9.2	7.8
Melanoma* SK-MEL-5	2.6	4.2	Colon ca. SW480	25.0	32.3
Squamous cell carcinoma SCC-4	9.0	8.4	Colon ca.* (SW480 met) SW620	5.0	4.6
Testis Pool	1.3	1.1	Colon ca. HT29	26.8	30.6
Prostate ca.* (bone met) PC-3	21.0	24.8	Colon ca. HCT-116	4.9	5.8
Prostate Pool	0.9	0.8	Colon ca. CaCo-2	13.6	10.4
Placenta	0.9	1.3	Colon cancer tissue	10.2	10.0
Uterus Pool	0.4	0.4	Colon ca. SW1116	5.1	3.6
Ovarian ca. OVCAR-3	77.4	66.9	Colon ca. Colo-205	1.8	1.5
Ovarian ca. SK-OV-3	42.0	36.3	Colon ca. SW-48	1.2	0.7
Ovarian ca. OVCAR-4	9.5	12.7	Colon Pool	1.9	1.3
Ovarian ca. OVCAR-5	39.2	44.4	Small Intestine Pool	0.6	1.0
Ovarian ca. IGROV-1	22.1	27.7	Stomach Pool	1.5	1.2
Ovarian ca. OVCAR-8	18.0	14.9	Bone Marrow Pool	0.6	0.5

Ovary	1.5	1.9	Fetal Heart	1.4	1.0
Breast ca. MCF-7	7.7	9.7	Heart Pool	0.7	0.8
Breast ca. MDA-MB-231	22.7	31.2	Lymph Node Pool	2.1	2.0
Breast ca. BT 549	13.2	10.1	Fetal Skeletal Muscle	0.9	0.5
Breast ca. T47D	100.0	100.0	Skeletal Muscle Pool	0.4	0.5
Breast ca. MDA-N	4.8	4.2	Spleen Pool	0.7	0.7
Breast Pool	1.6	1.6	Thymus Pool	1.8	2.2
Trachea	2.8	2.6	CNS cancer (glio/astro) U87-MG	5.9	6.0
Lung	0.2	0.1	CNS cancer (glio/astro) U-118-MG	11.7	11.2
Fetal Lung	11.4	8.3	CNS cancer (neuro;met) SK-N-AS	1.2	0.9
Lung ca. NCI-N417	1.4	0.7	CNS cancer (astro) SF-539	6.7	5.0
Lung ca. LX-1	10.5	11.0	CNS cancer (astro) SNB-75	22.8	32.3
Lung ca. NCI-H146	0.1	0.1	CNS cancer (glio) SNB-19	25.0	20.2
Lung ca. SHP-77	1.1	0.8	CNS cancer (glio) SF-295	35.1	38.2
Lung ca. A549	15.6	10.4	Brain (Amygdala) Pool	1.7	1.3
Lung ca. NCI-H526	5.4	1.6	Brain (cerebellum)	1.4	1.0
Lung ca. NCI-H23	18.9	20.6	Brain (fetal)	7.0	2.8
Lung ca. NCI-H460	11.5	9.3	Brain (Hippocampus) Pool	1.6	0.9
Lung ca. HOP-62	23.7	23.0	Cerebral Cortex Pool	1.9	0.9
Lung ca. NCI-H522	1.8	2.3	Brain (Substantia nigra) Pool	2.8	1.7
Liver	0.5	0.6	Brain (Thalamus) Pool	2.7	1.6
Fetal Liver	0.8	1.4	Brain (whole)	3.4	1.1
Liver ca. HepG2	15.5	12.6	Spinal Cord Pool	1.8	1.4
Kidney Pool	1.5	2.5	Adrenal Gland	0.2	0.4
Fetal Kidney	5.8	4.6	Pituitary gland Pool	0.3	0.2
Renal ca. 786-0	46.3	39.5	Salivary Gland	1.1	1.3
Renal ca. A498	13.8	7.9	Thyroid (female)	3.3	3.7
Renal ca. ACHN	14.3	15.9	Pancreatic ca. CAPAN2	23.7	27.7

Renal ca. UO-31	41.5	38.7	Pancreas Pool	3.0	4.1
-----------------	------	------	---------------	-----	-----

Table ACE. Panel 2.2

Tissue Name	Rel. Exp.(%) Ag3605, Run 173764229	Tissue Name	Rel. Exp.(%) Ag3605, Run 173764229
Normal Colon	4.7	Kidney Margin (OD04348)	100.0
Colon cancer (OD06064)	6.3	Kidney malignant cancer (OD06204B)	21.6
Colon Margin (OD06064)	2.3	Kidney normal adjacent tissue (OD06204E)	23.2
Colon cancer (OD06159)	2.1	Kidney Cancer (OD04450-01)	53.2
Colon Margin (OD06159)	1.8	Kidney Margin (OD04450-03)	21.3
Colon cancer (OD06297-04)	2.0	Kidney Cancer 8120613	1.1
Colon Margin (OD06297-05)	3.0	Kidney Margin 8120614	14.1
CC Gr.2 ascend colon (ODO3921)	3.6	Kidney Cancer 9010320	20.3
CC Margin (ODO3921)	1.3	Kidney Margin 9010321	15.0
Colon cancer metastasis (OD06104)	1.1	Kidney Cancer 8120607	71.7
Lung Margin (OD06104)	1.0	Kidney Margin 8120608	12.2
Colon mets to lung (OD04451-01)	4.5	Normal Uterus	7.3
Lung Margin (OD04451-02)	6.7	Uterine Cancer 064011	6.9
Normal Prostate	2.1	Normal Thyroid	4.3
Prostate Cancer (OD04410)	3.9	Thyroid Cancer 064010	27.0
Prostate Margin (OD04410)	3.4	Thyroid Cancer A302152	19.1
Normal Ovary	7.6	Thyroid Margin A302153	8.1
Ovarian cancer (OD06283-03)	27.9	Normal Breast	14.0
Ovarian Margin (OD06283-07)	2.0	Breast Cancer (OD04566)	13.4
Ovarian Cancer 064008	16.0	Breast Cancer 1024	35.1
Ovarian cancer (OD06145)	10.1	Breast Cancer (OD04590-01)	31.6
Ovarian Margin (OD06145)	8.2	Breast Cancer Mets (OD04590-03)	8.7
Ovarian cancer (OD06455-03)	28.9	Breast Cancer Metastasis (OD04655-05)	13.3
Ovarian Margin	1.9	Breast Cancer 064006	21.5

(OD06455-07)			
Normal Lung	2.9	Breast Cancer 9100266	17.7
Invasive poor diff. lung adeno (ODO4945-01)	9.4	Breast Margin 9100265	16.5
Lung Margin (ODO4945-03)	7.9	Breast Cancer A209073	13.2
Lung Malignant Cancer (OD03126)	7.5	Breast Margin A2090734	35.4
Lung Margin (OD03126)	7.0	Breast cancer (OD06083)	24.5
Lung Cancer (OD05014A)	17.0	Breast cancer node metastasis (OD06083)	21.5
Lung Margin (OD05014B)	11.7	Normal Liver	5.0
Lung cancer (OD06081)	12.2	Liver Cancer 1026	15.5
Lung Margin (OD06081)	2.4	Liver Cancer 1025	12.2
Lung Cancer (OD04237-01)	1.8	Liver Cancer 6004-T	7.8
Lung Margin (OD04237-02)	16.2	Liver Tissue 6004-N	6.1
Ocular Melanoma Metastasis	8.4	Liver Cancer 6005-T	25.0
Ocular Melanoma Margin (Liver)	2.9	Liver Tissue 6005-N	12.4
Melanoma Metastasis	4.0	Liver Cancer 064003	12.9
Melanoma Margin (Lung)	3.8	Normal Bladder	14.2
Normal Kidney	10.9	Bladder Cancer 1023	9.5
Kidney Ca, Nuclear grade 2 (OD04338)	35.4	Bladder Cancer A302173	12.2
Kidney Margin (OD04338)	20.4	Normal Stomach	8.7
Kidney Ca Nuclear grade 1/2 (OD04339)	52.9	Gastric Cancer 9060397	8.9
Kidney Margin (OD04339)	16.6	Stomach Margin 9060396	7.4
Kidney Ca, Clear cell type (OD04340)	16.6	Gastric Cancer 9060395	7.0
Kidney Margin (OD04340)	7.4	Stomach Margin 9060394	7.5
Kidney Ca, Nuclear grade 3 (OD04348)	11.2	Gastric Cancer 064005	6.9

Table ACF. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3605, Run 169943454	Rel. Exp.(%) Ag3974, Run 170739806	Tissue Name	Rel. Exp.(%) Ag3605, Run 169943454	Rel. Exp.(%) Ag3974, Run 170739806
-------------	--	--	-------------	--	--

Secondary Th1 act	1.0	1.2	HUVEC IL-1beta	15.6	18.9
Secondary Th2 act	5.1	8.0	HUVEC IFN gamma	12.9	16.7
Secondary Tr1 act	2.5	3.5	HUVEC TNF alpha + IFN gamma	37.6	34.9
Secondary Th1 rest	0.0	0.7	HUVEC TNF alpha + IL4	31.4	31.4
Secondary Th2 rest	0.4	0.2	HUVEC IL-11	14.9	13.9
Secondary Tr1 rest	0.4	1.2	Lung Microvascular EC none	79.0	100.0
Primary Th1 act	3.6	3.2	Lung Microvascular EC TNFalpha + IL-1beta	100.0	97.9
Primary Th2 act	1.1	2.0	Microvascular Dermal EC none	49.7	48.3
Primary Tr1 act	3.4	2.9	Microvascular Dermal EC TNFalpha + IL-1beta	56.6	47.0
Primary Th1 rest	0.9	0.4	Bronchial epithelium TNFalpha + IL1 beta	78.5	90.1
Primary Th2 rest	0.5	0.2	Small airway epithelium none	31.0	32.5
Primary Tr1 rest	0.2	0.3	Small airway epithelium TNFalpha + IL-1beta	81.8	93.3
CD45RA CD4 lymphocyte act	43.5	22.7	Coronary artery SMC rest	17.2	28.5
CD45RO CD4 lymphocyte act	5.0	5.5	Coronary artery SMC TNFalpha + IL-1beta	22.2	28.7
CD8 lymphocyte act	3.9	3.3	Astrocytes rest	80.1	55.1
Secondary CD8 lymphocyte rest	3.7	3.3	Astrocytes TNFalpha + IL-1beta	82.9	66.4
Secondary CD8 lymphocyte act	3.3	3.5	KU-812 (Basophil) rest	2.6	1.9
CD4 lymphocyte none	0.3	0.1	KU-812 (Basophil) PMA/ionomycin	0.6	2.8
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.3	0.4	CCD1106 (Keratinocytes) none	70.7	82.4
LAK cells rest	5.0	6.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	77.4	72.7
LAK cells IL-2	2.8	1.7	Liver cirrhosis	13.2	14.4
LAK cells IL-2+IL-12	1.4	1.8	NCI-H292 none	57.8	54.0
LAK cells IL-2+IFN gamma	2.3	1.1	NCI-H292 IL-4	62.4	78.5

LAK cells IL-2+ IL-18	2.9	1.6	NCI-H292 IL-9	61.6	79.6
LAK cells PMA/ionomycin	6.8	4.6	NCI-H292 IL-13	53.6	59.9
NK Cells IL-2 rest	1.6	1.9	NCI-H292 IFN gamma	67.4	71.7
Two Way MLR 3 day	10.7	12.4	HPAEC none	21.5	21.3
Two Way MLR 5 day	6.5	5.3	HPAEC TNF alpha + IL-1 beta	37.6	45.4
Two Way MLR 7 day	4.3	4.0	Lung fibroblast none	22.4	29.3
PBMC rest	0.0	0.6	Lung fibroblast TNF alpha + IL-1 beta	71.2	87.7
PBMC PWM	5.0	4.9	Lung fibroblast IL-4	16.2	23.3
PBMC PHA-L	5.5	3.4	Lung fibroblast IL-9	31.9	30.4
Ramos (B cell) none	0.4	0.4	Lung fibroblast IL-13	18.7	36.6
Ramos (B cell) ionomycin	0.2	0.2	Lung fibroblast IFN gamma	23.0	29.7
B lymphocytes PWM	2.4	3.0	Dermal fibroblast CCD1070 rest	15.9	27.2
B lymphocytes CD40L and IL-4	1.5	3.7	Dermal fibroblast CCD1070 TNF alpha	15.9	20.6
EOL-1 dbcAMP	3.4	3.1	Dermal fibroblast CCD1070 IL-1 beta	17.2	22.4
EOL-1 dbcAMP PMA/ionomycin	18.3	8.0	Dermal fibroblast IFN gamma	7.2	10.3
Dendritic cells none	14.8	9.0	Dermal fibroblast IL-4	7.2	8.0
Dendritic cells LPS	48.3	32.8	Dermal Fibroblasts rest	4.3	6.3
Dendritic cells anti-CD40	9.7	8.8	Neutrophils TNFa+LPS	0.0	0.9
Monocytes rest	0.9	1.4	Neutrophils rest	0.2	1.0
Monocytes LPS	66.4	81.2	Colon	7.0	5.8
Macrophages rest	16.2	9.7	Lung	23.3	23.3
Macrophages LPS	54.7	43.8	Thymus	5.8	7.3
HUVEC none	9.3	12.6	Kidney	23.2	33.2
HUVEC starved	14.4	25.2			

Table ACG. Panel CNS_1

Tissue Name	Rel. Exp.(%) Ag3605, Run 171648697	Tissue Name	Rel. Exp.(%) Ag3605, Run 171648697
BA4 Control	23.8	BA17 PSP	21.9
BA4 Control2	44.4	BA17 PSP2	14.4
BA4 Alzheimer's2	7.5	Sub Nigra Control	37.9

BA4 Parkinson's	66.0	Sub Nigra Control2	31.0
BA4 Parkinson's2	80.7	Sub Nigra Alzheimer's2	26.4
BA4 Huntington's	23.5	Sub Nigra Parkinson's2	80.1
BA4 Huntington's2	49.3	Sub Nigra Huntington's	76.3
BA4 PSP	19.5	Sub Nigra Huntington's2	29.7
BA4 PSP2	34.9	Sub Nigra PSP2	11.1
BA4 Depression	20.6	Sub Nigra Depression	34.4
BA4 Depression2	21.3	Sub Nigra Depression2	18.8
BA7 Control	53.2	Glob Palladus Control	40.3
BA7 Control2	47.6	Glob Palladus Control2	35.8
BA7 Alzheimer's	13.6	Glob Palladus Alzheimer's	20.0
BA7 Parkinson's	39.8	Glob Palladus Alzheimer's2	21.8
BA7 Parkinson's2	60.7	Glob Palladus Parkinson's	100.0
BA7 Huntington's	41.8	Glob Palladus Parkinson's2	25.0
BA7 Huntington's2	62.9	Glob Palladus PSP	17.9
BA7 PSP	36.1	Glob Palladus PSP2	7.2
BA7 PSP2	25.0	Glob Palladus Depression	15.5
BA7 Depression	20.0	Temp Pole Control	15.3
BA9 Control	36.6	Temp Pole Control2	76.8
BA9 Control2	83.5	Temp Pole Alzheimer's	14.3
BA9 Alzheimer's	17.1	Temp Pole Alzheimer's2	14.7
BA9 Alzheimer's2	34.4	Temp Pole Parkinson's	76.3
BA9 Parkinson's	70.7	Temp Pole Parkinson's2	77.4
BA9 Parkinson's2	74.2	Temp Pole Huntington's	39.8
BA9 Huntington's	55.5	Temp Pole PSP	7.4
BA9 Huntington's2	45.1	Temp Pole PSP2	8.1
BA9 PSP	28.7	Temp Pole Depression2	31.6
BA9 PSP2	8.2	Cing Gyr Control	82.4
BA9 Depression	18.0	Cing Gyr Control2	82.4
BA9 Depression2	0.0	Cing Gyr Alzheimer's	27.4
BA17 Control	74.7	Cing Gyr Alzheimer's2	36.3
BA17 Control2	86.5	Cing Gyr Parkinson's	46.3
BA17 Alzheimer's	20.3	Cing Gyr Parkinson's2	42.6
BA17 Parkinson's	75.3	Cing Gyr Huntington's	70.7
BA17 Parkinson's2	85.3	Cing Gyr Huntington's2	37.6
BA17 Huntington's	47.0	Cing Gyr PSP	21.6
BA17 Huntington's2	26.6	Cing Gyr PSP2	13.9
BA17 Depression	24.8	Cing Gyr Depression	21.3

BA17 Depression2	41.2	Cing Gyr Depression2	32.5
------------------	------	----------------------	------

Table ACH. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3605, Run 260268655	Tissue Name	Rel. Exp.(%) Ag3605, Run 260268655
Colon cancer 1	11.6	Bladder cancer NAT 2	0.2
Colon cancer NAT 1	5.3	Bladder cancer NAT 3	1.7
Colon cancer 2	18.6	Bladder cancer NAT 4	1.8
Colon cancer NAT 2	3.2	Adenocarcinoma of the prostate 1	2.4
Colon cancer 3	22.2	Adenocarcinoma of the prostate 2	2.6
Colon cancer NAT 3	7.7	Adenocarcinoma of the prostate 3	5.2
Colon malignant cancer 4	23.5	Adenocarcinoma of the prostate 4	21.9
Colon normal adjacent tissue 4	2.0	Prostate cancer NAT 5	5.4
Lung cancer 1	57.0	Adenocarcinoma of the prostate 6	4.4
Lung NAT 1	3.1	Adenocarcinoma of the prostate 7	3.8
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	1.8
Lung NAT 2	3.9	Adenocarcinoma of the prostate 9	10.7
Squamous cell carcinoma 3	41.2	Prostate cancer NAT 10	0.7
Lung NAT 3	1.7	Kidney cancer 1	38.7
metastatic melanoma 1	5.4	Kidney NAT 1	20.6
Melanoma 2	3.2	Kidney cancer 2	66.4
Melanoma 3	3.7	Kidney NAT 2	39.8
metastatic melanoma 4	12.5	Kidney cancer 3	64.6
metastatic melanoma 5	7.9	Kidney NAT 3	19.6
Bladder cancer 1	0.5	Kidney cancer 4	16.4
Bladder cancer NAT 1	0.0	Kidney NAT 4	15.6
Bladder cancer 2	6.0		

CNS_neurodegeneration_v1.0 Summary: Ag3605/Ag3974 Two experiments with two different probe and primer sets confirm the expression of this gene at moderate levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and

those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag3605/Ag3974 Two experiments with the same probe and primer set produce results that are in excellent agreement. The expression of the CG94946-01 gene appears to be highest in a sample derived from a breast cancer cell line (T47D) (CTs=22.5-25.3). In addition, there appears to be substantial expression in other samples derived from breast cancer cell lines, ovarian cancer cell lines, kidney cancer cell lines, lung cancer cell lines, colon cancer cell lines and brain cancer cell lines. Thus, the expression of this gene could be used to distinguish T47D cells from other samples in the panel. Moreover, therapeutic modulation of this gene, through the use of small molecule drugs, protein therapeutics, or antibodies could be of benefit in the treatment of breast, ovarian, kidney, lung, colon or brain cancer.

Among metabolic tissues, this gene has low-to-moderate levels of expression in adrenal, pituitary, adult and fetal heart, adult and fetal liver, adult and fetal skeletal muscle, and adipose. This gene product has high levels of expression (CT values = 27) in pancreas and thyroid. Thus, this gene product may be important for the pathogenesis, diagnosis, and/or treatment of metabolic and endocrine diseases, including obesity, Types 1 and 2 diabetes and thyroidopathies. In support of this hypothesis, decreased glomerular expression of agrin has been observed in diabetic nephropathy (Yard BA, Exp Nephrol 2001;9(3):214-22).

In addition, this gene is expressed at moderate levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. The CG94946-01 gene encodes a protein with homology to agrin, a neuronal aggregating factor that induces the aggregation of acetylcholine receptors and other postsynaptic proteins on muscle fibers and is crucial for the formation of the neuromuscular junction. Agrin also plays an important role in defining neuronal responses to excitatory neurotransmitters both in vitro and in vivo (Hilgenberg LG, Mol Cell Neurosci 2002 Jan;19(1):97-110 and Bixby JL, J Neurobiol 2002 Feb 5;50(2):164-79). The CG59841-01 gene expression in the central nervous system is consistent with the hypothesis that this protein may have similar functions as agrin. Therefore, this gene may play a role in central nervous system disorders such as

Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Agrin has also been implicated in the formation of senile plaques in Alzheimer's disease and in the acetylcholine synapse/neuromuscular junction (van Horssen J, Acta Neuropathol (Berl) 2001 Dec;102(6):604-14). In addition, an agrin minigene rescued dystrophic symptoms in a mouse model of muscular dystrophy (Moll J, Nature. 2001 Sep 20;413(6853):302-7). Therefore, this gene product may be used as a treatment or cure for congenital muscular dystrophies. Furthermore, this gene product is also an excellent drug target in AD or in any disease involving the neuromuscular junction or the acetylcholine system.

Panel 2.2 Summary: Ag3605 Expression of the CG94946-01 gene is highest in a sample of normal kidney (CT = 27.4). In addition, expression of this gene appears to be upregulated in a number of ovarian and renal cancers when compared to the matched control margins. Thus, expression of this gene could be used as a marker for ovarian and renal carcinoma. Furthermore, therapeutic modulation of the activity of this gene or its protein product, using small molecule drugs, antibodies or protein therapeutics, could be of benefit in the treatment of renal and ovarian cancer.

Panel 4.1D Summary: Ag3605/Ag3974 Two experiments with two different probe and primer sets produce results that are in very good agreement. Highest expression of the CG94946-01 gene is highest in lung microvascular endothelial cells (CTs=27.3-28.5), microvascular dermal endothelial cells, mucoepidermoid cell line NCI-H292, astrocytes, and keratinocytes. This gene encodes a protein with homology to agrin. Recently, it has been demonstrated that agrin, an aggregating protein crucial for formation of the neuromuscular junction, is also important for T cell signaling in the immune system (Khan AA, Science 2001 Jun 1;292(5522):1681-6). In addition, agrin has been identified as a potential disease target for autoimmune disorders at the neuromuscular junction, including multiple sclerosis (Liyanage Y, Muscle Nerve 2002 Jan;25(1):4-16). Therefore, small molecule drug, antibody or protein therapeutics designed against the protein encoded by the CG59841-01 gene could reduce or inhibit inflammation in asthma, emphysema, allergy, psoriasis, muscular dystrophy and multiple sclerosis.

Panel CNS_1 Summary: Ag3605 This panel confirms the expression of the CG94946-01 gene at low levels in the brains of an independent group of individuals.

Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders

- General oncology screening panel_v_2.4 Summary:** Ag3605 Highest expression of the CG94946-01 gene is seen in lung cancer (CT=26.8). In addition, higher levels of expression are seen in lung and kidney cancers when compared to expression in normal adjacent tissue. Thus, expression of this gene could be used as a marker of lung and kidney cancers. Furthermore, therapeutic modulation of the expression or function of this gene product may be useful in the treatment of lung and kidney cancers.

10 AD. CG95165-01: Adenylate Cyclase, Type II

Expression of gene CG95165-01 was assessed using the primer-probe set Ag3991, described in Table ADA. Results of the RTQ-PCR runs are shown in Tables ADB, ADC and ADD.

Table ADA. Probe Name Ag3991

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aacgtggcatcctgaagagt-3'	20	3261	317
Probe	TET-5'-caccttcattttggcaagaagactgt-3'-TAMRA	26	3281	318
Reverse	5'-gttgcagtcagaaagtgtgtga-3'	22	3322	319

15

Table ADB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag3991, Run 212349022	Tissue Name	Rel. Exp.(%) Ag3991, Run 212349022
AD 1 Hippo	27.9	Control (Path) 3 Temporal Ctx	15.9
AD 2 Hippo	46.0	Control (Path) 4 Temporal Ctx	35.4
AD 3 Hippo	14.9	AD 1 Occipital Ctx	26.2
AD 4 Hippo	13.8	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	61.6	AD 3 Occipital Ctx	14.9
AD 6 Hippo	97.9	AD 4 Occipital Ctx	31.2
Control 2 Hippo	43.8	AD 5 Occipital Ctx	64.6
Control 4 Hippo	26.1	AD 6 Occipital Ctx	26.4
Control (Path) 3 Hippo	15.9	Control 1 Occipital Ctx	9.6
AD 1 Temporal Ctx	39.0	Control 2 Occipital Ctx	71.2
AD 2 Temporal Ctx	55.9	Control 3 Occipital Ctx	21.9
AD 3 Temporal Ctx	15.0	Control 4 Occipital Ctx	16.5
AD 4 Temporal Ctx	33.2	Control (Path) 1 Occipital Ctx	81.2

AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	9.3
AD 5 Sup Temporal Ctx	60.7	Control (Path) 3 Occipital Ctx	7.4
AD 6 Inf Temporal Ctx	74.7	Control (Path) 4 Occipital Ctx	9.6
AD 6 Sup Temporal Ctx	90.1	Control 1 Parietal Ctx	15.0
Control 1 Temporal Ctx	15.1	Control 2 Parietal Ctx	56.3
Control 2 Temporal Ctx	64.2	Control 3 Parietal Ctx	23.2
Control 3 Temporal Ctx	25.0	Control (Path) 1 Parietal Ctx	99.3
Control 3 Temporal Ctx	15.5	Control (Path) 2 Parietal Ctx	28.3
Control (Path) 1 Temporal Ctx	65.1	Control (Path) 3 Parietal Ctx	11.6
Control (Path) 2 Temporal Ctx	43.5	Control (Path) 4 Parietal Ctx	44.4

Table ADC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag3991, Run 170739808	Tissue Name	Rel. Exp.(%) Ag3991, Run 170739808
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	1.5	HUVEC TNF alpha + IFN gamma	4.5
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	37.4

Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	15.8
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	2.8
LAK cells IL-2+IL-12	0.0	NCI-H292 none	1.8
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	2.8
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	8.2
Dendritic cells none	0.0	Dermal fibroblast IL-4	42.3
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	39.5
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	2.0
Monocytes LPS	0.0	Colon	17.7

Macrophages rest	2.4	Lung	8.8
Macrophages LPS	0.0	Thymus	15.2
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

Table ADD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3991, Run 268168218	Tissue Name	Rel. Exp.(%) Ag3991, Run 268168218
Colon cancer 1	0.2	Bladder cancer NAT 2	0.1
Colon cancer NAT 1	1.1	Bladder cancer NAT 3	0.0
Colon cancer 2	0.3	Bladder cancer NAT 4	0.9
Colon cancer NAT 2	0.5	Adenocarcinoma of the prostate 1	7.0
Colon cancer 3	0.6	Adenocarcinoma of the prostate 2	0.9
Colon cancer NAT 3	1.4	Adenocarcinoma of the prostate 3	4.8
Colon malignant cancer 4	0.1	Adenocarcinoma of the prostate 4	0.8
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	2.0
Lung cancer 1	0.1	Adenocarcinoma of the prostate 6	1.0
Lung NAT 1	0.1	Adenocarcinoma of the prostate 7	1.4
Lung cancer 2	0.2	Adenocarcinoma of the prostate 8	0.9
Lung NAT 2	0.1	Adenocarcinoma of the prostate 9	8.3
Squamous cell carcinoma 3	1.7	Prostate cancer NAT 10	0.5
Lung NAT 3	0.2	Kidney cancer 1	0.6
metastatic melanoma 1	3.8	Kidney NAT 1	0.4
Melanoma 2	2.3	Kidney cancer 2	100.0
Melanoma 3	2.0	Kidney NAT 2	0.0
metastatic melanoma 4	4.6	Kidney cancer 3	0.1
metastatic melanoma 5	1.6	Kidney NAT 3	0.1
Bladder cancer 1	0.3	Kidney cancer 4	0.6
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.4
Bladder cancer 2	1.2		

CNS_neurodegeneration_v1.0 Summary: Ag3991 This panel confirms the expression of the CG95165-01 gene at low levels in the brains of an independent group of individuals, but no differential expression of this gene was detected between Alzheimer's

5 diseased postmortem brains and those of non-demented controls in this experiment.

However, expression in the brain suggests that this gene may play a role in central nervous system disorders such as Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

The CG95165-01 gene codes for a homologue of rat adenylyl cyclase. In
5 Drosophila, mutations in adenylyl cyclase and other members of cAMP pathway have been shown to reduce the ability of fruit flies to learn or to remember (Waddell S and Quinn WG, 2001, Trends Genet 17(12):719-26, PMID: 11718926). In addition, derangement of second messenger, adenylyl cyclase, system closely parallels ischemic neuronal damage and persistent enhancement of this cAMP signaling pathway is important
10 for neuronal survival in acute cerebral ischemia (Tanaka K, 2001, Prog Neurobiol 65(2):173-207, PMID: 11403878). Therefore, therapeutic modulation of adenylyl cyclase encoded by this gene may be useful in the treatment of learning disorders and acute cerebral ischemia.

General_screening_panel_v1.4 Summary: Ag3991 Results from one
15 experiment with the CG95165-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 4.1D Summary: Ag3991 Highest expression of the CG95165-01 gene is detected in kidney (CT=32). In addition, moderate to low expression of this gene is seen in thymus, colon, dermal fibroblast and astrocytes. Thus, expression of this gene can be used
20 to distinguish these samples from other samples used in this panel. In addition, therapeutic modulation of this gene may be beneficial in the treatment of autoimmune and inflammatory diseases that affect brain, colon and kidney, including inflammatory bowel diseases, lupus and glomerulonephritis.

General oncology screening panel_v_2.4 Summary: Ag3991 Highest
25 expression of the CG95165-01 gene is detected in kidney cancer (CT=27.8). Interestingly, expression of this gene is higher in cancer sample compared to the adjacent control sample (CT=40). In addition, moderate expression of this gene is seen in prostate adenocarcinoma and melanoma. This expression is low/undetectable in the normal samples used in this panel. Therefore, therapeutic modulation of this gene product through the use of small
30 molecule drug or antibodies could be beneficial in the treatment of melanoma, prostate and kidney cancers.

AE. CG95175-01: Ephrin Type-A Receptor 7 Precursor

Expression of gene CG95175-01 was assessed using the primer-probe sets Ag3992 and Ag612, described in Tables AEA and AEB. Results of the RTQ-PCR runs are shown in Table AEC.

5

Table AEA. Probe Name Ag3992

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - accactatggtgaggctacaga - 3'	22	2427	320
Probe	TET-5' - ctatgggcccgcctcccagacact - 3' - TAMRA	23	2466	321
Reverse	5' - agagctgaagtggccaaact - 3'	20	2491	322

Table AEB. Probe Name Ag612

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - gccgctcccagacactt - 3'	18	2472	323
Probe	TET-5' - ccacttcagctctgccagtgcgtg - 3' - TAMRA	25	2498	324
Reverse	5' - cccacatgatgatgccgaa - 3'	19	2529	325

Table AEC. Panel 4D

Tissue Name	Rel. Exp.(%) Ag612, Run 145645058	Tissue Name	Rel. Exp.(%) Ag612, Run 145645058
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	12.2	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	4.6
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0

CD45RO CD4 lymphocyte act	4.6	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	5.3	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	5.3	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	9.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	1.0
LAK cells IL-2	0.0	Liver cirrhosis	61.6
LAK cells IL-2+IL-12	0.0	Lupus kidney	0.0
LAK cells IL-2+IFN gamma	5.1	NCI-H292 none	32.5
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-4	46.7
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-9	58.6
NK Cells IL-2 rest	0.0	NCI-H292 IL-13	89.5
Two Way MLR 3 day	0.0	NCI-H292 IFN gamma	100.0
Two Way MLR 5 day	0.0	HPAEC none	0.0
Two Way MLR 7 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
PBMC rest	0.0	Lung fibroblast none	0.0
PBMC PWM	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PHA-L	8.9	Lung fibroblast IL-4	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IL-13	9.0
B lymphocytes PWM	8.5	Lung fibroblast IFN gamma	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 rest	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
Dendritic cells none	5.3	Dermal fibroblast IFN gamma	0.0
Dendritic cells LPS	4.5	Dermal fibroblast IL-4	0.0
Dendritic cells anti-CD40	0.0	IBD Colitis 2	10.7

Monocytes rest	0.0	IBD Crohn's	0.0
Monocytes LPS	1.2	Colon	51.8
Macrophages rest	0.0	Lung	24.7
Macrophages LPS	0.0	Thymus	4.1
HUVEC none	0.0	Kidney	4.6
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag3992 Expression of the CG95175-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag3992 Expression of the CG95175-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag3992 Expression of the CG95175-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4D Summary: Ag612 Highest expression of the CG95175-01 gene is detected in IFN gamma treated NCI-H292 cells (CT=33). Moderate to low expression of this gene is also seen in cytokine treated and untreated NCI-H292 cells, liver cirrhosis and colon tissue samples. Therefore, expression of this gene can be used to distinguish these samples from other samples used in this panel. In addition, therapeutic modulation of this gene can be used for the treatment of chronic obstructive pulmonary disease, asthma, allergy, and emphysema, liver cirrhosis, autoimmune and inflammatory disease affecting colon including Crohn's disease and ulcerative colitis.

Panel CNS_1 Summary: Ag3992 Expression of the CG95175-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag3992 Expression of the CG95175-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

AF. CG95693-01: DPY-19 Protein I Like Protein

Expression of gene CG95693-01 was assessed using the primer-probe set Ag4026, described in Table AFA. Results of the RTQ-PCR runs are shown in Tables AFB, AFC, AFD and AFE.

Table AFA. Probe Name Ag4026

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tggcattctaacagtgatgtca-3'	22	54	326
Probe	TET-5'-tgcaaacctctgtaatcaatggagca-3'-TAMRA	26	87	327
Reverse	5'-aaagttcttctgaggcaaatc-3'	22	130	328

Table AFB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4026, Run 212391720	Tissue Name	Rel. Exp.(%) Ag4026, Run 212391720
AD 1 Hippo	7.6	Control (Path) 3 Temporal Ctx	2.0
AD 2 Hippo	24.0	Control (Path) 4 Temporal Ctx	27.7
AD 3 Hippo	12.4	AD 1 Occipital Ctx	15.3
AD 4 Hippo	9.3	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	77.4	AD 3 Occipital Ctx	4.8
AD 6 Hippo	38.7	AD 4 Occipital Ctx	24.5
Control 2 Hippo	11.2	AD 5 Occipital Ctx	23.8
Control 4 Hippo	23.8	AD 6 Occipital Ctx	30.6
Control (Path) 3 Hippo	12.8	Control 1 Occipital Ctx	1.6
AD 1 Temporal Ctx	12.2	Control 2 Occipital Ctx	47.0
AD 2 Temporal Ctx	26.4	Control 3 Occipital Ctx	13.2
AD 3 Temporal Ctx	5.6	Control 4 Occipital Ctx	7.9
AD 4 Temporal Ctx	12.5	Control (Path) 1 Occipital Ctx	52.1
AD 5 Inf Temporal Ctx	84.7	Control (Path) 2 Occipital Ctx	13.4
AD 5 Sup Temporal Ctx	52.5	Control (Path) 3 Occipital Ctx	4.0
AD 6 Inf Temporal Ctx	55.5	Control (Path) 4 Occipital Ctx	14.7
AD 6 Sup Temporal Ctx	100.0	Control 1 Parietal Ctx	5.1
Control 1 Temporal Ctx	7.9	Control 2 Parietal Ctx	47.6
Control 2 Temporal Ctx	18.0	Control 3 Parietal Ctx	6.6
Control 3 Temporal Ctx	12.9	Control (Path) 1 Parietal Ctx	51.1
Control 4 Temporal Ctx	6.0	Control (Path) 2 Parietal Ctx	17.9
Control (Path) 1 Temporal Ctx	47.0	Control (Path) 3 Parietal Ctx	4.8
Control (Path) 2 Temporal Ctx	22.8	Control (Path) 4 Parietal Ctx	33.4

Table AFC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4026, Run 218425757	Tissue Name	Rel. Exp.(%) Ag4026, Run 218425757
Adipose	0.9	Renal ca. TK-10	12.7

Melanoma* Hs688(A).T	0.6	Bladder	7.6
Melanoma* Hs688(B).T	0.9	Gastric ca. (liver met.) NCI-N87	14.1
Melanoma* M14	2.5	Gastric ca. KATO III	31.0
Melanoma* LOXIMVI	5.4	Colon ca. SW-948	4.3
Melanoma* SK-MEL-5	16.8	Colon ca. SW480	49.0
Squamous cell carcinoma SCC-4	0.8	Colon ca.* (SW480 met) SW620	25.9
Testis Pool	91.4	Colon ca. HT29	1.6
Prostate ca.* (bone met) PC-3	5.6	Colon ca. HCT-116	0.9
Prostate Pool	2.6	Colon ca. CaCo-2	6.3
Placenta	0.4	Colon cancer tissue	20.4
Uterus Pool	0.8	Colon ca. SW1116	0.9
Ovarian ca. OVCAR-3	3.3	Colon ca. Colo-205	6.2
Ovarian ca. SK-OV-3	31.2	Colon ca. SW-48	6.5
Ovarian ca. OVCAR-4	2.0	Colon Pool	5.5
Ovarian ca. OVCAR-5	10.2	Small Intestine Pool	7.2
Ovarian ca. IGROV-1	11.6	Stomach Pool	3.8
Ovarian ca. OVCAR-8	3.8	Bone Marrow Pool	0.0
Ovary	2.5	Fetal Heart	5.2
Breast ca. MCF-7	12.4	Heart Pool	2.1
Breast ca. MDA-MB-231	10.9	Lymph Node Pool	7.7
Breast ca. BT 549	23.7	Fetal Skeletal Muscle	1.9
Breast ca. T47D	10.2	Skeletal Muscle Pool	2.0
Breast ca. MDA-N	7.2	Spleen Pool	6.2
Breast Pool	7.6	Thymus Pool	7.6
Trachea	8.4	CNS cancer (glio/astro) U87-MG	7.6
Lung	5.2	CNS cancer (glio/astro) U-118-MG	27.9
Fetal Lung	15.6	CNS cancer (neuro;met) SK-N-AS	6.8
Lung ca. NCI-N417	0.4	CNS cancer (astro) SF-539	10.7
Lung ca. LX-1	29.9	CNS cancer (astro) SNB-75	60.7
Lung ca. NCI-H146	1.8	CNS cancer (glio) SNB-19	14.9
Lung ca. SHP-77	1.6	CNS cancer (glio) SF-295	100.0
Lung ca. A549	0.9	Brain (Amygdala) Pool	4.8
Lung ca. NCI-H526	2.1	Brain (cerebellum)	0.4
Lung ca. NCI-H23	52.9	Brain (fetal)	3.8
Lung ca. NCI-H460	3.8	Brain (Hippocampus) Pool	6.2

Lung ca. HOP-62	2.4	Cerebral Cortex Pool	6.0
Lung ca. NCI-H522	4.2	Brain (Substantia nigra) Pool	6.6
Liver	0.0	Brain (Thalamus) Pool	8.8
Fetal Liver	14.8	Brain (whole)	2.9
Liver ca. HepG2	1.8	Spinal Cord Pool	9.3
Kidney Pool	7.9	Adrenal Gland	1.6
Fetal Kidney	10.6	Pituitary gland Pool	3.4
Renal ca. 786-0	14.4	Salivary Gland	0.7
Renal ca. A498	3.3	Thyroid (female)	2.5
Renal ca. ACHN	3.1	Pancreatic ca. CAPAN2	9.0
Renal ca. UO-31	2.9	Pancreas Pool	5.1

Table AFD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4026, Run 171613290	Tissue Name	Rel. Exp.(%) Ag4026, Run 171613290
Secondary Th1 act	1.3	HUVEC IL-1beta	0.7
Secondary Th2 act	4.2	HUVEC IFN gamma	0.9
Secondary Tr1 act	2.0	HUVEC TNF alpha + IFN gamma	1.2
Secondary Th1 rest	2.9	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	3.3	HUVEC IL-11	0.0
Secondary Tr1 rest	2.6	Lung Microvascular EC none	1.0
Primary Th1 act	3.4	Lung Microvascular EC TNFalpha + IL-1beta	1.0
Primary Th2 act	9.2	Microvascular Dermal EC none	0.0
Primary Tr1 act	3.9	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	1.1	Bronchial epithelium TNFalpha + IL1beta	0.5
Primary Th2 rest	0.5	Small airway epithelium none	0.0
Primary Tr1 rest	2.0	Small airway epithelium TNFalpha + IL-1beta	1.5
CD45RA CD4 lymphocyte act	7.3	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	7.7	Coronary artery SMC TNFalpha + IL-1beta	0.4
CD8 lymphocyte act	8.1	Astrocytes rest	2.0
Secondary CD8 lymphocyte rest	2.9	Astrocytes TNFalpha + IL-1beta	1.4
Secondary CD8 lymphocyte act	2.7	KU-812 (Basophil) rest	74.2

CD4 lymphocyte none	1.3	KU-812 (Basophil) PMA/ionomycin	100.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	7.6	CCD1106 (Keratinocytes) none	3.7
LAK cells rest	3.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	2.8
LAK cells IL-2	10.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	2.2	NCI-H292 none	3.2
LAK cells IL-2+IFN gamma	5.8	NCI-H292 IL-4	4.0
LAK cells IL-2+ IL-18	5.7	NCI-H292 IL-9	3.0
LAK cells PMA/ionomycin	1.4	NCI-H292 IL-13	3.0
NK Cells IL-2 rest	12.9	NCI-H292 IFN gamma	2.5
Two Way MLR 3 day	4.8	HPAEC none	0.5
Two Way MLR 5 day	2.1	HPAEC TNF alpha + IL-1 beta	1.8
Two Way MLR 7 day	3.1	Lung fibroblast none	0.0
PBMC rest	0.5	Lung fibroblast TNF alpha + IL-1 beta	1.4
PBMC PWM	5.6	Lung fibroblast IL-4	0.4
PBMC PHA-L	5.3	Lung fibroblast IL-9	1.0
Ramos (B cell) none	23.0	Lung fibroblast IL-13	2.2
Ramos (B cell) ionomycin	24.0	Lung fibroblast IFN gamma	2.3
B lymphocytes PWM	3.3	Dermal fibroblast CCD1070 rest	3.3
B lymphocytes CD40L and IL-4	12.2	Dermal fibroblast CCD1070 TNF alpha	12.9
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	1.4
Dendritic cells none	1.0	Dermal fibroblast IL-4	1.2
Dendritic cells LPS	0.8	Dermal Fibroblasts rest	3.7
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.6	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	11.7
Macrophages rest	4.0	Lung	6.6
Macrophages LPS	0.0	Thymus	22.4
HUVEC none	0.9	Kidney	49.3
HUVEC starved	1.7		

Table AFE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4026, Run	Tissue Name	Rel. Exp.(%) Ag4026, Run
-------------	-----------------------------	-------------	-----------------------------

	268362901		268362901
Colon cancer 1	23.8	Bladder cancer NAT 2	0.6
Colon NAT 1	0.7	Bladder cancer NAT 3	0.0
Colon cancer 2	3.9	Bladder cancer NAT 4	2.2
Colon cancer NAT 2	2.0	Adenocarcinoma of the prostate 1	30.8
Colon cancer 3	63.3	Adenocarcinoma of the prostate 2	1.3
Colon cancer NAT 3	4.9	Adenocarcinoma of the prostate 3	3.5
Colon malignant cancer 4	16.8	Adenocarcinoma of the prostate 4	38.7
Colon normal adjacent tissue 4	1.2	Prostate cancer NAT 5	0.4
Lung cancer 1	13.9	Adenocarcinoma of the prostate 6	1.6
Lung NAT 1	0.9	Adenocarcinoma of the prostate 7	0.4
Lung cancer 2	16.3	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.3	Adenocarcinoma of the prostate 9	6.1
Squamous cell carcinoma 3	4.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.5	Kidney cancer 1	12.8
metastatic melanoma 1	30.1	Kidney NAT 1	5.0
Melanoma 2	0.0	Kidney cancer 2	100.0
Melanoma 3	1.5	Kidney NAT 2	6.3
metastatic melanoma 4	12.4	Kidney cancer 3	52.9
metastatic melanoma 5	26.8	Kidney NAT 3	4.2
Bladder cancer 1	0.9	Kidney cancer 4	12.1
Bladder cancer NAT 1	0.0	Kidney NAT 4	2.6
Bladder cancer 2	1.0		

CNS_neurodegeneration_v1.0 Summary: Ag4026 This panel confirms the expression of the CG95693-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4026 Highest expression of the CG95693-01 gene is detected in CNS cancer SF-295 cell line (CT=28.3). Significant expression of this gene is associated with cluster of cancer cell lines including CNS, colon, renal, lung, breast, ovarian, pancreatic, melanoma and prostate cancer cell lines. Therefore, therapeutic modulation of this gene could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, fetal liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, this gene is expressed at much higher levels in fetal (CT=31) when compared to adult liver (CT=40). This observation suggests that expression of this gene can be used to distinguish fetal from adult liver. In addition, the relative overexpression of this gene in fetal liver suggests that the protein product may enhance liver growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of liver related diseases.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4026 Highest expression of the CG95693-01 gene is detected in PMA/ionomycin treated basophils (CT=31). In addition, this gene is expressed to a lesser extent in untreated KU-812 cells. Therefore, antibody or small molecule therapies designed with the protein encoded for by this gene could block or inhibit inflammation or tissue damage due to basophil activation in response to asthma, allergies, hypersensitivity reactions, psoriasis, and viral infections.

Moderate to low levels of expression of this gene is also detected in Ramos B cells, activated lymphocytes, IL-2 treated LAK cells, TNF alpha treated dermal fibroblasts, and normal tissues represented by colon, lung, thymus and kidney. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4026 In this panel the highest expression of the CG95693-01 gene is detected in a kidney cancer sample (CT=31.4). Interestingly, expression of this gene is higher in kidney, prostate, and lung cancer samples as compared to adjacent control samples as well as the melanoma samples in this panel. Therapeutic modulation of this gene, through the use of small molecule drugs, protein therapeutics or antibodies could be of benefit in the treatment of melanomas as well as kidney, prostate and lung cancers.

AG. CG95814-01: Rho-GTPase-Activating Protein 4

Expression of gene CG95814-01 was assessed using the primer-probe set Ag4031, described in Table AGA. Results of the RTQ-PCR runs are shown in Tables AGB, AGC, AGD, AGE and AGF.

Table AGA. Probe Name Ag4031

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aaccaatgcatctgtcttcaag-3'	22	1053	329
Probe	TET-5'-tccatgacctatctgaccttattgatca-3'-TAMRA	28	1082	330
Reverse	5'-tgcattgtagcctaagtcaca-3'	21	1114	331

Table AGB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4031, Run 212396115	Tissue Name	Rel. Exp.(%) Ag4031, Run 212396115
AD 1 Hippo	18.2	Control (Path) 3 Temporal Ctx	10.4
AD 2 Hippo	42.0	Control (Path) 4 Temporal Ctx	29.3
AD 3 Hippo	13.2	AD 1 Occipital Ctx	25.9
AD 4 Hippo	11.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	96.6	AD 3 Occipital Ctx	7.7
AD 6 Hippo	83.5	AD 4 Occipital Ctx	19.9
Control 2 Hippo	42.9	AD 5 Occipital Ctx	21.8
Control 4 Hippo	20.4	AD 6 Occipital Ctx	55.9
Control (Path) 3 Hippo	5.0	Control 1 Occipital Ctx	6.5
AD 1 Temporal Ctx	26.2	Control 2 Occipital Ctx	68.3
AD 2 Temporal Ctx	29.1	Control 3 Occipital Ctx	15.1
AD 3 Temporal Ctx	8.9	Control 4 Occipital Ctx	10.9
AD 4 Temporal Ctx	0.0	Control (Path) 1 Occipital Ctx	88.3
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	6.5

AD 5 SupTemporal Ctx	61.1	Control (Path) 3 Occipital Ctx	6.6
AD 6 Inf Temporal Ctx	71.2	Control (Path) 4 Occipital Ctx	20.9
AD 6 Sup Temporal Ctx	66.9	Control 1 Parietal Ctx	7.4
Control 1 Temporal Ctx	5.8	Control 2 Parietal Ctx	48.6
Control 2 Temporal Ctx	47.6	Control 3 Parietal Ctx	13.4
Control 3 Temporal Ctx	12.9	Control (Path) 1 Parietal Ctx	68.3
Control 4 Temporal Ctx	9.9	Control (Path) 2 Parietal Ctx	15.4
Control (Path) 1 Temporal Ctx	61.1	Control (Path) 3 Parietal Ctx	6.7
Control (Path) 2 Temporal Ctx	26.2	Control (Path) 4 Parietal Ctx	45.7

Table AGC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4031, Run 218525121	Tissue Name	Rel. Exp.(%) Ag4031, Run 218525121
Adipose	10.7	Renal ca. TK-10	34.9
Melanoma* Hs688(A).T	28.1	Bladder	13.4
Melanoma* Hs688(B).T	24.8	Gastric ca. (liver met.) NCI-N87	23.0
Melanoma* M14	50.3	Gastric ca. KATO III	27.9
Melanoma* LOXIMVI	25.9	Colon ca. SW-948	1.9
Melanoma* SK-MEL-5	66.0	Colon ca. SW480	66.0
Squamous cell carcinoma SCC-4	14.3	Colon ca.* (SW480 met) SW620	14.1
Testis Pool	4.5	Colon ca. HT29	6.1
Prostate ca.* (bone met) PC-3	34.4	Colon ca. HCT-116	26.8
Prostate Pool	3.1	Colon ca. CaCo-2	6.8
Placenta	15.6	Colon cancer tissue	8.0
Uterus Pool	4.5	Colon ca. SW1116	1.4
Ovarian ca. OVCAR-3	15.8	Colon ca. Colo-205	3.5
Ovarian ca. SK-OV-3	34.6	Colon ca. SW-48	3.4
Ovarian ca. OVCAR-4	15.7	Colon Pool	12.9
Ovarian ca. OVCAR-5	25.3	Small Intestine Pool	9.6
Ovarian ca. IGROV-1	28.9	Stomach Pool	5.8
Ovarian ca. OVCAR-8	14.7	Bone Marrow Pool	3.2
Ovary	9.8	Fetal Heart	8.2
Breast ca. MCF-7	17.1	Heart Pool	3.9
Breast ca. MDA-MB-231	55.1	Lymph Node Pool	8.4

Breast ca. BT 549	42.6	Fetal Skeletal Muscle	12.3
Breast ca. T47D	43.8	Skeletal Muscle Pool	3.7
Breast ca. MDA-N	26.8	Spleen Pool	16.8
Breast Pool	10.2	Thymus Pool	8.8
Trachea	6.7	CNS cancer (glio/astro) U87-MG	37.4
Lung	3.1	CNS cancer (glio/astro) U-118-MG	48.0
Fetal Lung	37.4	CNS cancer (neuro;met) SK-N-AS	30.1
Lung ca. NCI-N417	4.6	CNS cancer (astro) SF-539	45.1
Lung ca. LX-1	15.5	CNS cancer (astro) SNB-75	83.5
Lung ca. NCI-H146	7.0	CNS cancer (glio) SNB-19	32.5
Lung ca. SHP-77	73.2	CNS cancer (glio) SF-295	46.7
Lung ca. A549	20.7	Brain (Amygdala) Pool	14.5
Lung ca. NCI-H526	4.2	Brain (cerebellum)	100.0
Lung ca. NCI-H23	37.4	Brain (fetal)	55.9
Lung ca. NCI-H460	20.0	Brain (Hippocampus) Pool	16.3
Lung ca. HOP-62	24.1	Cerebral Cortex Pool	15.5
Lung ca. NCI-H522	31.0	Brain (Substantia nigra) Pool	18.0
Liver	0.9	Brain (Thalamus) Pool	23.7
Fetal Liver	12.6	Brain (whole)	32.1
Liver ca. HepG2	15.4	Spinal Cord Pool	18.9
Kidney Pool	18.0	Adrenal Gland	9.3
Fetal Kidney	13.7	Pituitary gland Pool	2.2
Renal ca. 786-0	54.0	Salivary Gland	1.1
Renal ca. A498	8.0	Thyroid (female)	2.1
Renal ca. ACHN	19.1	Pancreatic ca. CAPAN2	31.9
Renal ca. UO-31	24.5	Pancreas Pool	8.5

Table AGD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4031, Run 171616912	Tissue Name	Rel. Exp.(%) Ag4031, Run 171616912
Secondary Th1 act	49.3	HUVEC IL-1beta	49.7
Secondary Th2 act	100.0	HUVEC IFN gamma	43.5
Secondary Tr1 act	62.0	HUVEC TNF alpha + IFN gamma	32.5
Secondary Th1 rest	13.2	HUVEC TNF alpha + IL4	28.3
Secondary Th2 rest	13.9	HUVEC IL-11	25.2

Secondary Tr1 rest	11.9	Lung Microvascular EC none	55.9
Primary Th1 act	14.2	Lung Microvascular EC TNFalpha + IL-1beta	31.9
Primary Th2 act	23.0	Microvascular Dermal EC none	42.6
Primary Tr1 act	22.5	Microvascular Dermal EC TNFalpha + IL-1beta	27.5
Primary Th1 rest	13.6	Bronchial epithelium TNFalpha + IL1beta	19.8
Primary Th2 rest	9.2	Small airway epithelium none	6.9
Primary Tr1 rest	16.3	Small airway epithelium TNFalpha + IL-1beta	13.3
CD45RA CD4 lymphocyte act	26.8	Coronary artery SMC rest	22.2
CD45RO CD4 lymphocyte act	22.7	Coronary artery SMC TNFalpha + IL-1beta	24.7
CD8 lymphocyte act	19.3	Astrocytes rest	38.4
Secondary CD8 lymphocyte rest	14.4	Astrocytes TNFalpha + IL-1beta	26.8
Secondary CD8 lymphocyte act	21.5	KU-812 (Basophil) rest	9.3
CD4 lymphocyte none	4.9	KU-812 (Basophil) PMA/ionomycin	49.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	21.9	CCD1106 (Keratinocytes) none	38.2
LAK cells rest	24.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	33.2
LAK cells IL-2	19.3	Liver cirrhosis	5.0
LAK cells IL-2+IL-12	12.4	NCI-H292 none	20.6
LAK cells IL-2+IFN gamma	17.0	NCI-H292 IL-4	24.5
LAK cells IL-2+ IL-18	22.8	NCI-H292 IL-9	37.1
LAK cells PMA/ionomycin	8.4	NCI-H292 IL-13	25.2
NK Cells IL-2 rest	30.1	NCI-H292 IFN gamma	28.9
Two Way MLR 3 day	22.2	HPAEC none	34.9
Two Way MLR 5 day	19.5	HPAEC TNF alpha + IL-1 beta	43.8
Two Way MLR 7 day	22.5	Lung fibroblast none	26.6
PBMC rest	6.6	Lung fibroblast TNF alpha + IL-1 beta	24.7
PBMC PWM	13.7	Lung fibroblast IL-4	13.6
PBMC PHA-L	19.6	Lung fibroblast IL-9	23.0
Ramos (B cell) none	20.2	Lung fibroblast IL-13	14.9

Ramos (B cell) ionomycin	23.7	Lung fibroblast IFN gamma	24.8
B lymphocytes PWM	22.7	Dermal fibroblast CCD1070 rest	37.9
B lymphocytes CD40L and IL-4	22.7	Dermal fibroblast CCD1070 TNF alpha	54.7
EOL-1 dbcAMP	25.9	Dermal fibroblast CCD1070 IL-1 beta	21.5
EOL-1 dbcAMP PMA/ionomycin	27.2	Dermal fibroblast IFN gamma	29.7
Dendritic cells none	36.1	Dermal fibroblast IL-4	46.3
Dendritic cells LPS	76.8	Dermal Fibroblasts rest	26.4
Dendritic cells anti-CD40	32.5	Neutrophils TNFa+LPS	2.6
Monocytes rest	33.4	Neutrophils rest	9.7
Monocytes LPS	42.0	Colon	4.7
Macrophages rest	18.6	Lung	27.0
Macrophages LPS	26.8	Thymus	20.6
HUVEC none	37.1	Kidney	16.6
HUVEC starved	45.1		

Table AGE. Panel CNS_1

Tissue Name	Rel. Exp.(%) Ag4031, Run 180912028	Tissue Name	Rel. Exp.(%) Ag4031, Run 180912028
BA4 Control	24.1	BA17 PSP	30.1
BA4 Control2	50.7	BA17 PSP2	7.1
BA4 Alzheimer's2	3.5	Sub Nigra Control	37.9
BA4 Parkinson's	68.3	Sub Nigra Control2	34.4
BA4 Parkinson's2	75.8	Sub Nigra Alzheimer's2	25.5
BA4 Huntington's	42.3	Sub Nigra Parkinson's2	77.9
BA4 Huntington's2	6.5	Sub Nigra Huntington's	100.0
BA4 PSP	8.2	Sub Nigra Huntington's2	69.7
BA4 PSP2	30.4	Sub Nigra PSP2	12.8
BA4 Depression	49.7	Sub Nigra Depression	20.7
BA4 Depression2	15.5	Sub Nigra Depression2	14.9
BA7 Control	31.9	Glob Palladus Control	20.2
BA7 Control2	31.6	Glob Palladus Control2	15.1
BA7 Alzheimer's2	7.4	Glob Palladus Alzheimer's	18.6
BA7 Parkinson's	27.5	Glob Palladus Alzheimer's2	8.3
BA7 Parkinson's2	47.6	Glob Palladus Parkinson's	81.8
BA7 Huntington's	75.3	Glob Palladus Parkinson's2	19.9

BA7 Huntington's2	48.6	Glob Palladus PSP	9.5
BA7 PSP	38.7	Glob Palladus PSP2	8.0
BA7 PSP2	23.3	Glob Palladus Depression	10.2
BA7 Depression	13.6	Temp Pole Control	12.4
BA9 Control	39.0	Temp Pole Control2	40.1
BA9 Control2	83.5	Temp Pole Alzheimer's	5.5
BA9 Alzheimer's	5.0	Temp Pole Alzheimer's2	6.1
BA9 Alzheimer's2	19.8	Temp Pole Parkinson's	35.1
BA9 Parkinson's	42.6	Temp Pole Parkinson's2	38.4
BA9 Parkinson's2	50.3	Temp Pole Huntington's	49.0
BA9 Huntington's	59.9	Temp Pole PSP	3.3
BA9 Huntington's2	11.8	Temp Pole PSP2	5.3
BA9 PSP	15.4	Temp Pole Depression2	11.2
BA9 PSP2	4.4	Cing Gyr Control	57.8
BA9 Depression	10.7	Cing Gyr Control2	32.1
BA9 Depression2	8.2	Cing Gyr Alzheimer's	20.9
BA17 Control	44.8	Cing Gyr Alzheimer's2	10.6
BA17 Control2	46.0	Cing Gyr Parkinson's	38.7
BA17 Alzheimer's2	12.6	Cing Gyr Parkinson's2	46.0
BA17 Parkinson's	40.9	Cing Gyr Huntington's	76.3
BA17 Parkinson's2	57.8	Cing Gyr Huntington's2	38.2
BA17 Huntington's	44.4	Cing Gyr PSP	26.2
BA17 Huntington's2	24.7	Cing Gyr PSP2	9.5
BA17 Depression	12.7	Cing Gyr Depression	10.2
BA17 Depression2	32.5	Cing Gyr Depression2	19.6

Table AGF. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4031, Run 268362913	Tissue Name	Rel. Exp.(%) Ag4031, Run 268362913
Colon cancer 1	13.4	Bladder cancer NAT 2	1.0
Colon NAT 1	2.2	Bladder cancer NAT 3	1.3
Colon cancer 2	27.4	Bladder cancer NAT 4	3.8
Colon cancer NAT 2	5.6	Adenocarcinoma of the prostate 1	9.3
Colon cancer 3	15.2	Adenocarcinoma of the prostate 2	2.1
Colon cancer NAT 3	15.8	Adenocarcinoma of the prostate 3	9.0
Colon malignant cancer 4	35.4	Adenocarcinoma of the	8.8

		prostate 4	
Colon normal adjacent tissue 4	2.7	Prostate cancer NAT 5	6.3
Lung cancer 1	13.9	Adenocarcinoma of the prostate 6	3.1
Lung NAT 1	4.3	Adenocarcinoma of the prostate 7	6.1
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	1.2
Lung NAT 2	9.3	Adenocarcinoma of the prostate 9	25.2
Squamous cell carcinoma 3	29.5	Prostate cancer NAT 10	1.7
Lung NAT 3	1.6	Kidney cancer 1	34.2
metastatic melanoma 1	20.2	Kidney NAT 1	19.1
Melanoma 2	8.5	Kidney cancer 2	42.0
Melanoma 3	8.0	Kidney NAT 2	38.4
metastatic melanoma 4	93.3	Kidney cancer 3	33.4
metastatic melanoma 5	65.1	Kidney NAT 3	19.1
Bladder cancer 1	1.6	Kidney cancer 4	15.5
Bladder cancer NAT 1	0.0	Kidney NAT 4	8.4
Bladder cancer 2	6.0		

CNS_neurodegeneration_v1.0 Summary: Ag4031 This panel confirms the expression of the CG95814-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between

5 Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4031 Highest expression of the CG95814-01 gene is detected in brain (cerebellum) sample (CT=25.6). High expression of

10 this gene is seen in all the region of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

15 The CG95814-01 gene codes for a homolog of RHO-GTPase activating protein (RHOGAP). RHOGAP stimulates the GTPase activity of Rho protein. Since Rho proteins

have been implicated in ischemia (Trapp et al., 2001, Mol Cell Neurosci 17(5):883-94, PMID: 11358485), inhibitors for the RHOGAP encoded by this gene may have a therapeutic utility in the treatment of pathological aspects following brain damage including neuronal death.

5 High expression of this gene is also observed in cluster of cancer cell lines including pancreatic, CNS, colon, renal, lung, breast, ovarian, prostate, squamous cell carcinoma, and melanoma cancer cell lines. Therefore, therapeutic modulation of this gene can be useful in treatment of these cancers.

10 Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

15 Interestingly, this gene is expressed at much higher levels in fetal (CT=27-28) when compared to adult lung and liver (CTs=30-32). This observation suggests that expression of this gene can be used to distinguish fetal from adult lung and liver. In addition, the relative overexpression of this gene in fetal tissue suggests that the protein product may enhance growth or development of lung and liver in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the
20 protein encoded by this gene could be useful in treatment of liver and lung related diseases.

Panel 4.1D Summary: Ag4031 Highest expression of the CG95814-01 gene is detected in activated secondary Th2 cells (CT=27). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in
25 health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and
30 tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional

therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

- 5 **Panel CNS_1 Summary:** Ag4031 This panel confirms the expression of the CG95814-01 gene at low levels in the brains of an independent group of individuals. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

- 10 **General oncology screening panel_v_2.4 Summary:** Ag4031 Highest expression of the CG95814-01 gene is detected in lung cancer sample (CT=27). Higher expression of this gene is also seen in melanoma, colon, kidney, prostate cancers and this expression is lower in corresponding control samples. Thus, expression of this gene can be used as diagnostic marker and therapeutic modulation of this gene product may be useful in the treatment of these cancers.

15

AH. CG95824-01: Rho GAP

Expression of gene CG95824-01 was assessed using the primer-probe set Ag4032, described in Table AHA. Results of the RTQ-PCR runs are shown in Tables AHB, AHC, AHD and AHE.

20

Table AHA. Probe Name Ag4032

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctcctacctggaaggctacatc-3'	22	2475	332
Probe	TET-5'-atggaccccgtaaatgacactaaagg-3'-TAMRA	26	2506	333
Reverse	5'-cagtacttggcatacgtgctta-3'	22	2540	334

Table AHB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4032, Run 212396260	Tissue Name	Rel. Exp.(%) Ag4032, Run 212396260
AD 1 Hippo	9.9	Control (Path) 3 Temporal Ctx	5.9
AD 2 Hippo	14.8	Control (Path) 4 Temporal Ctx	27.0
AD 3 Hippo	5.2	AD 1 Occipital Ctx	14.9
AD 4 Hippo	3.5	AD 2 Occipital Ctx (Missing)	0.0

AD 5 Hippo	100.0	AD 3 Occipital Ctx	6.2
AD 6 Hippo	29.3	AD 4 Occipital Ctx	18.2
Control 2 Hippo	22.8	AD 5 Occipital Ctx	45.4
Control 4 Hippo	2.5	AD 6 Occipital Ctx	16.0
Control (Path) 3 Hippo	3.5	Control 1 Occipital Ctx	1.4
AD 1 Temporal Ctx	10.5	Control 2 Occipital Ctx	84.7
AD 2 Temporal Ctx	19.2	Control 3 Occipital Ctx	21.8
AD 3 Temporal Ctx	4.7	Control 4 Occipital Ctx	1.6
AD 4 Temporal Ctx	11.7	Control (Path) 1 Occipital Ctx	53.2
AD 5 Inf Temporal Ctx	73.2	Control (Path) 2 Occipital Ctx	14.2
AD 5 Sup Temporal Ctx	27.0	Control (Path) 3 Occipital Ctx	3.3
AD 6 Inf Temporal Ctx	32.3	Control (Path) 4 Occipital Ctx	17.2
AD 6 Sup Temporal Ctx	39.2	Control 1 Parietal Ctx	5.6
Control 1 Temporal Ctx	4.2	Control 2 Parietal Ctx	32.8
Control 2 Temporal Ctx	35.6	Control 3 Parietal Ctx	23.0
Control 3 Temporal Ctx	15.5	Control (Path) 1 Parietal Ctx	62.0
Control 3 Temporal Ctx	6.7	Control (Path) 2 Parietal Ctx	18.8
Control (Path) 1 Temporal Ctx	41.8	Control (Path) 3 Parietal Ctx	3.0
Control (Path) 2 Temporal Ctx	32.8	Control (Path) 4 Parietal Ctx	37.4

Table AHC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4032, Run 218525131	Tissue Name	Rel. Exp.(%) Ag4032, Run 218525131
Adipose	0.5	Renal ca. TK-10	22.7
Melanoma* Hs688(A).T	2.7	Bladder	4.2
Melanoma* Hs688(B).T	2.4	Gastric ca. (liver met.) NCI-N87	31.4
Melanoma* M14	4.4	Gastric ca. KATO III	41.8
Melanoma* LOXIMVI	12.4	Colon ca. SW-948	8.6
Melanoma* SK-MEL-5	25.3	Colon ca. SW480	17.2
Squamous cell carcinoma SCC-4	5.0	Colon ca.* (SW480 met) SW620	11.3
Testis Pool	19.6	Colon ca. HT29	5.8
Prostate ca.* (bone met) PC-3	30.8	Colon ca. HCT-116	40.9
Prostate Pool	2.0	Colon ca. CaCo-2	9.7
Placenta	2.5	Colon cancer tissue	8.8
Uterus Pool	0.0	Colon ca. SW1116	2.5

Ovarian ca. OVCAR-3	20.7	Colon ca. Colo-205	2.3
Ovarian ca. SK-OV-3	16.5	Colon ca. SW-48	5.2
Ovarian ca. OVCAR-4	10.1	Colon Pool	2.1
Ovarian ca. OVCAR-5	35.1	Small Intestine Pool	1.1
Ovarian ca. IGROV-1	4.2	Stomach Pool	1.6
Ovarian ca. OVCAR-8	15.8	Bone Marrow Pool	0.5
Ovary	2.6	Fetal Heart	7.3
Breast ca. MCF-7	37.9	Heart Pool	1.5
Breast ca. MDA-MB-231	18.6	Lymph Node Pool	2.3
Breast ca. BT 549	84.7	Fetal Skeletal Muscle	5.9
Breast ca. T47D	58.2	Skeletal Muscle Pool	0.3
Breast ca. MDA-N	3.0	Spleen Pool	1.7
Breast Pool	3.0	Thymus Pool	3.9
Trachea	10.2	CNS cancer (glio/astro) U87-MG	9.9
Lung	0.1	CNS cancer (glio/astro) U-118-MG	10.4
Fetal Lung	17.9	CNS cancer (neuro:met) SK-N-AS	20.9
Lung ca. NCI-N417	3.2	CNS cancer (astro) SF-539	5.0
Lung ca. LX-1	9.9	CNS cancer (astro) SNB-75	12.6
Lung ca. NCI-H146	10.4	CNS cancer (glio) SNB-19	12.0
Lung ca. SHP-77	50.0	CNS cancer (glio) SF-295	11.2
Lung ca. A549	16.8	Brain (Amygdala) Pool	33.0
Lung ca. NCI-H526	18.4	Brain (cerebellum)	87.7
Lung ca. NCI-H23	5.3	Brain (fetal)	100.0
Lung ca. NCI-H460	5.5	Brain (Hippocampus) Pool	33.4
Lung ca. HOP-62	5.1	Cerebral Cortex Pool	69.3
Lung ca. NCI-H522	33.7	Brain (Substantia nigra) Pool	55.5
Liver	0.2	Brain (Thalamus) Pool	51.1
Fetal Liver	1.6	Brain (whole)	69.3
Liver ca. HepG2	4.8	Spinal Cord Pool	23.2
Kidney Pool	2.0	Adrenal Gland	4.9
Fetal Kidney	15.8	Pituitary gland Pool	2.7
Renal ca. 786-0	4.5	Salivary Gland	2.5
Renal ca. A498	1.6	Thyroid (female)	3.1
Renal ca. ACHN	4.9	Pancreatic ca. CAPAN2	10.2
Renal ca. UO-31	6.5	Pancreas Pool	6.8

Table AHD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4032, Run	Tissue Name	Rel. Exp.(%) Ag4032, Run
-------------	-----------------------------	-------------	-----------------------------

	171616913		171616913
Secondary Th1 act	0.0	HUVEC IL-1beta	11.1
Secondary Th2 act	1.4	HUVEC IFN gamma	24.7
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	7.7
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	17.9
Secondary Th2 rest	1.1	HUVEC IL-11	14.3
Secondary Tr1 rest	0.0	Lung Microvascular EC none	11.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	20.7
Primary Th2 act	1.3	Microvascular Dermal EC none	20.2
Primary Tr1 act	1.0	Microvascular Dermal EC TNFalpha + IL-1beta	10.2
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	19.5
Primary Th2 rest	0.0	Small airway epithelium none	6.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	10.5
CD45RA CD4 lymphocyte act	15.4	Coronary artery SMC rest	15.8
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	6.7
CD8 lymphocyte act	0.0	Astrocytes rest	31.0
Secondary CD8 lymphocyte rest	6.4	Astrocytes TNFalpha + IL-1beta	22.2
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	13.7
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	14.5
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	40.9
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	22.1
LAK cells IL-2	0.0	Liver cirrhosis	8.8
LAK cells IL-2+IL-12	0.0	NCI-H292 none	40.3
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	56.6
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	51.8
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	31.9
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	100.0
Two Way MLR 3 day	0.0	HPAEC none	19.2
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	32.3

Two Way MLR 7 day	0.0	Lung fibroblast none	35.1
PBMC rest	1.3	Lung fibroblast TNF alpha + IL-1 beta	13.1
PBMC PWM	0.0	Lung fibroblast IL-4	17.8
PBMC PHA-L	0.0	Lung fibroblast IL-9	34.6
Ramos (B cell) none	10.8	Lung fibroblast IL-13	36.6
Ramos (B cell) ionomycin	12.9	Lung fibroblast IFN gamma	31.2
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	24.5
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	15.7
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	12.6
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	11.0
Dendritic cells none	3.5	Dermal fibroblast IL-4	40.6
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	25.2
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	13.9
Macrophages rest	0.0	Lung	6.0
Macrophages LPS	0.0	Thymus	15.5
HUVEC none	15.2	Kidney	26.6
HUVEC starved	10.7		

Table AHE. Panel CNS_1

Tissue Name	Rel. Exp.(%) Ag4032, Run 180912030	Tissue Name	Rel. Exp.(%) Ag4032, Run 180912030
BA4 Control	19.1	BA17 PSP	31.6
BA4 Control2	36.9	BA17 PSP2	17.1
BA4 Alzheimer's2	13.0	Sub Nigra Control	10.1
BA4 Parkinson's	56.6	Sub Nigra Control2	24.3
BA4 Parkinson's2	88.9	Sub Nigra Alzheimer's2	3.8
BA4 Huntington's	69.3	Sub Nigra Parkinson's2	21.0
BA4 Huntington's2	11.9	Sub Nigra Huntington's	23.8
BA4 PSP	13.8	Sub Nigra Huntington's2	19.3
BA4 PSP2	29.7	Sub Nigra PSP2	6.6
BA4 Depression	14.0	Sub Nigra Depression	3.0
BA4 Depression2	13.1	Sub Nigra Depression2	2.8
BA7 Control	71.7	Glob Palladus Control	0.9

BA7 Control2	39.8	Glob Palladus Control2	0.0
BA7 Alzheimer's2	15.0	Glob Palladus Alzheimer's	2.5
BA7 Parkinson's	38.4	Glob Palladus Alzheimer's2	2.9
BA7 Parkinson's2	64.6	Glob Palladus Parkinson's	76.3
BA7 Huntington's	53.2	Glob Palladus Parkinson's2	3.5
BA7 Huntington's2	42.9	Glob Palladus PSP	0.0
BA7 PSP	44.1	Glob Palladus PSP2	3.2
BA7 PSP2	37.6	Glob Palladus Depression	0.0
BA7 Depression	8.8	Temp Pole Control	10.2
BA9 Control	17.7	Temp Pole Control2	36.1
BA9 Control2	100.0	Temp Pole Alzheimer's	8.7
BA9 Alzheimer's	7.1	Temp Pole Alzheimer's2	6.4
BA9 Alzheimer's2	25.9	Temp Pole Parkinson's	32.8
BA9 Parkinson's	65.1	Temp Pole Parkinson's2	18.7
BA9 Parkinson's2	49.3	Temp Pole Huntington's	31.0
BA9 Huntington's	41.5	Temp Pole PSP	8.1
BA9 Huntington's2	34.9	Temp Pole PSP2	3.8
BA9 PSP	11.4	Temp Pole Depression2	5.9
BA9 PSP2	3.6	Cing Gyr Control	37.4
BA9 Depression	18.8	Cing Gyr Control2	28.5
BA9 Depression2	13.7	Cing Gyr Alzheimer's	9.9
BA17 Control	92.7	Cing Gyr Alzheimer's2	7.2
BA17 Control2	57.8	Cing Gyr Parkinson's	26.2
BA17 Alzheimer's2	27.2	Cing Gyr Parkinson's2	8.5
BA17 Parkinson's	62.9	Cing Gyr Huntington's	43.5
BA17 Parkinson's2	74.7	Cing Gyr Huntington's2	5.6
BA17 Huntington's	77.4	Cing Gyr PSP	2.8
BA17 Huntington's2	26.1	Cing Gyr PSP2	1.4
BA17 Depression	18.4	Cing Gyr Depression	0.0
BA17 Depression2	39.2	Cing Gyr Depression2	10.8

CNS_neurodegeneration_v1.0 Summary: Ag4032 This panel confirms the expression of the CG95824-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between
5 Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4032 Highest expression of the CG95824-01 gene is detected in fetal brain (CT=29.8). High expression of this gene is seen all the regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord.

5 Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression and therapeutic modulation of this gene product may be beneficial in the treatment of these CNS disorders.

The CG95824-01 gene codes for a RHO-GTPase activating protein (RHOGAP).
10 RHOGAP stimulates the GTPase activity of Rho proteins. Since Rho proteins have been implicated in ischemia (Trapp et al., 2001, Mol Cell Neurosci 17(5):883-94, PMID: 11358485), inhibitors for the RHOGAP encoded by this gene may have a therapeutic utility in the treatment of pathological aspects following brain damage including neuronal death.

15 Moderate levels of expression of this gene is also seen in cluster of cancer cell lines including melanoma, CNS, colon, breast, ovarian, lung, gastric, renal, pancreatic and prostate cancer cell lines. Therefore, therapeutic modulation of this gene product through the use of small molecule target may be useful in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high
20 to moderate levels in pancreas, adrenal gland, thyroid, fetal heart, and fetal skeletal muscle. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, expression of this gene is higher in fetal (CTs=32-33) as compared to adult lung, heart and skeletal muscle (CTs>36). Therefore, expression of this gene can be
25 used to distinguish fetal from these adult tissue samples. In addition, the relative overexpression of this gene in fetal suggests that the protein product may enhance growth or development of these tissues in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the RHOGAP encoded by this gene could be useful in treatment of muscle, lung and heart related diseases.

30 **Panel 4.1D Summary:** Ag4032 Highest expression of the CG95824-01 gene is detected in IFN gamma treated NCI-H292 cells (CT=33.4). Moderate to low expression of this gene is seen in NCI-H292 cells, HPAEC, lung fibroblasts, IL-4 treated dermal

fibroblasts. Therefore, therapeutics designed with the protein encoded by this gene may reduce or eliminate symptoms caused by inflammation in lung epithelia, in chronic obstructive pulmonary disease, asthma, allergy, and emphysema.

- Panel CNS_1 Summary: Ag4032 This panel confirms the expression of the
 5 CG95824-01 gene at low levels in the brains of an independent group of individuals. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

AI. CG96231-01 and CG96231-02: OTU-Like Cysteine Protease

- 10 Expression of gene CG96231-01 and full length physical clone CG96231-02 was assessed using the primer-probe set Ag4043, described in Table AIA. Results of the RTQ-PCR runs are shown in Tables AIB, AIC, AID and AIE. Please note that CG96231-02 represents a full-length physical clone of the CG96231-01 gene, validating the prediction of the gene sequence.

15

Table AIA. Probe Name Ag-4043

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agcgtaacttcctgatcca-3'	20	815	335
Probe	TET-5'-cacctcctctgaccattttctcctct-3'-TAMRA	26	839	336
Reverse	5'-tcctctgctaattccagtgtt-3'	22	887	337

Table AIB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4043, Run 214151952	Tissue Name	Rel. Exp.(%) Ag4043, Run 214151952
AD 1 Hippo	13.8	Control (Path) 3 Temporal Ctx	4.5
AD 2 Hippo	15.4	Control (Path) 4 Temporal Ctx	28.3
AD 3 Hippo	8.6	AD 1 Occipital Ctx	40.6
AD 4 Hippo	8.1	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	98.6	AD 3 Occipital Ctx	10.2
AD 6 Hippo	72.2	AD 4 Occipital Ctx	19.8
Control 2 Hippo	20.6	AD 5 Occipital Ctx	41.2
Control 4 Hippo	8.2	AD 6 Occipital Ctx	31.4
Control (Path) 3 Hippo	6.3	Control 1 Occipital Ctx	3.3
AD 1 Temporal Ctx	16.6	Control 2 Occipital Ctx	43.8
AD 2 Temporal Ctx	19.3	Control 3 Occipital Ctx	11.1
AD 3 Temporal Ctx	8.5	Control 4 Occipital Ctx	7.4

AD 4 Temporal Ctx	20.6	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	74.2	Control (Path) 2 Occipital Ctx	8.7
AD 5 Sup Temporal Ctx	20.4	Control (Path) 3 Occipital Ctx	2.0
AD 6 Inf Temporal Ctx	87.7	Control (Path) 4 Occipital Ctx	12.9
AD 6 Sup Temporal Ctx	70.2	Control 1 Parietal Ctx	6.2
Control 1 Temporal Ctx	4.5	Control 2 Parietal Ctx	44.1
Control 2 Temporal Ctx	16.7	Control 3 Parietal Ctx	14.3
Control 3 Temporal Ctx	11.5	Control (Path) 1 Parietal Ctx	62.0
Control 4 Temporal Ctx	5.0	Control (Path) 2 Parietal Ctx	20.6
Control (Path) 1 Temporal Ctx	35.1	Control (Path) 3 Parietal Ctx	1.4
Control (Path) 2 Temporal Ctx	27.4	Control (Path) 4 Parietal Ctx	40.3

Table AIC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4043, Run 218426211	Tissue Name	Rel. Exp.(%) Ag4043, Run 218426211
Adipose	4.4	Renal ca. TK-10	19.8
Melanoma* Hs688(A).T	5.9	Bladder	10.4
Melanoma* Hs688(B).T	4.5	Gastric ca. (liver met.) NCI-N87	33.9
Melanoma* M14	8.8	Gastric ca. KATO III	27.9
Melanoma* LOXIMVI	15.6	Colon ca. SW-948	6.5
Melanoma* SK-MEL-5	19.1	Colon ca. SW480	27.2
Squamous cell carcinoma SCC-4	13.7	Colon ca.* (SW480 met) SW620	11.0
Testis Pool	4.9	Colon ca. HT29	10.5
Prostate ca.* (bone met) PC-3	46.3	Colon ca. HCT-116	27.7
Prostate Pool	2.2	Colon ca. CaCo-2	18.0
Placenta	3.5	Colon cancer tissue	6.5
Uterus Pool	2.3	Colon ca. SW1116	2.5
Ovarian ca. OVCAR-3	17.3	Colon ca. Colo-205	4.7
Ovarian ca. SK-OV-3	37.1	Colon ca. SW-48	4.0
Ovarian ca. OVCAR-4	3.7	Colon Pool	4.5
Ovarian ca. OVCAR-5	17.3	Small Intestine Pool	5.3
Ovarian ca. IGROV-1	9.7	Stomach Pool	5.5
Ovarian ca. OVCAR-8	5.0	Bone Marrow Pool	2.6
Ovary	4.3	Fetal Heart	5.3
Breast ca. MCF-7	0.7	Heart Pool	2.3

Breast ca. MDA-MB-231	29.9	Lymph Node Pool	6.7
Breast ca. BT 549	14.7	Fetal Skeletal Muscle	1.8
Breast ca. T47D	36.3	Skeletal Muscle Pool	4.4
Breast ca. MDA-N	4.1	Spleen Pool	7.9
Breast Pool	4.7	Thymus Pool	11.5
Trachea	6.7	CNS cancer (glio/astro) U87-MG	15.4
Lung	0.9	CNS cancer (glio/astro) U-118-MG	11.2
Fetal Lung	34.4	CNS cancer (neuro;met) SK-N-AS	16.3
Lung ca. NCI-N417	3.2	CNS cancer (astro) SF-539	5.8
Lung ca. LX-1	27.5	CNS cancer (astro) SNB-75	14.5
Lung ca. NCI-H146	5.0	CNS cancer (glio) SNB-19	8.9
Lung ca. SHP-77	12.2	CNS cancer (glio) SF-295	18.3
Lung ca. A549	14.3	Brain (Amygdala) Pool	4.5
Lung ca. NCI-H526	2.3	Brain (cerebellum)	4.2
Lung ca. NCI-H23	10.7	Brain (fetal)	8.1
Lung ca. NCI-H460	9.9	Brain (Hippocampus) Pool	5.5
Lung ca. HOP-62	6.8	Cerebral Cortex Pool	7.6
Lung ca. NCI-H522	7.3	Brain (Substantia nigra) Pool	8.0
Liver	0.5	Brain (Thalamus) Pool	8.8
Fetal Liver	100.0	Brain (whole)	9.3
Liver ca. HepG2	5.0	Spinal Cord Pool	8.3
Kidney Pool	7.6	Adrenal Gland	4.2
Fetal Kidney	7.9	Pituitary gland Pool	2.3
Renal ca. 786-0	15.9	Salivary Gland	1.6
Renal ca. A498	2.3	Thyroid (female)	3.5
Renal ca. ACHN	20.7	Pancreatic ca. CAPAN2	16.8
Renal ca. UO-31	8.3	Pancreas Pool	6.0

Table A1D. Panel 4.1D.

Tissue Name	Rel. Exp.(%) Ag4043, Run 171616960	Tissue Name	Rel. Exp.(%) Ag4043, Run 171616960
Secondary Th1 act	19.2	HUVEC IL-1beta	9.2
Secondary Th2 act	25.5	HUVEC IFN gamma	8.9
Secondary Tr1 act	14.3	HUVEC TNF alpha + IFN gamma	9.3
Secondary Th1 rest	5.2	HUVEC TNF alpha + IL4	5.3

Secondary Th2 rest	4.9	HUVEC IL-11	3.4
Secondary Tr1 rest	8.5	Lung Microvascular EC none	7.2
Primary Th1 act	10.7	Lung Microvascular EC TNFalpha + IL-1beta	3.5
Primary Th2 act	12.8	Microvascular Dermal EC none	5.0
Primary Tr1 act	16.6	Microvascular Dermal EC TNFalpha + IL-1beta	3.6
Primary Th1 rest	5.2	Bronchial epithelium TNFalpha + IL-1beta	6.4
Primary Th2 rest	2.5	Small airway epithelium none	3.2
Primary Tr1 rest	5.5	Small airway epithelium TNFalpha + IL-1beta	9.7
CD45RA CD4 lymphocyte act	11.1	Coronary artery SMC rest	3.0
CD45RO CD4 lymphocyte act	20.7	Coronary artery SMC TNFalpha + IL-1beta	4.9
CD8 lymphocyte act	12.8	Astrocytes rest	3.2
Secondary CD8 lymphocyte rest	11.4	Astrocytes TNFalpha + IL-1beta	3.3
Secondary CD8 lymphocyte act	8.0	KU-812 (Basophil) rest	51.4
CD4 lymphocyte none	6.0	KU-812 (Basophil) PMA/ionomycin	100.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	11.6	CCD1106 (Keratinocytes) none	13.6
LAK cells rest	10.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	6.7
LAK cells IL-2	14.9	Liver cirrhosis	5.6
LAK cells IL-2+IL-12	11.8	NCI-H292 none	13.9
LAK cells IL-2+IFN gamma	5.1	NCI-H292 IL-4	14.5
LAK cells IL-2+ IL-18	6.7	NCI-H292 IL-9	15.6
LAK cells PMA/ionomycin	10.7	NCI-H292 IL-13	14.0
NK Cells IL-2 rest	12.2	NCI-H292 IFN gamma	13.3
Two Way MLR 3 day	10.1	HPAEC none	3.1
Two Way MLR 5 day	8.6	HPAEC TNF alpha + IL-1 beta	7.4
Two Way MLR 7 day	8.2	Lung fibroblast none	3.7
PBMC rest	6.7	Lung fibroblast TNF alpha + IL- 1 beta	2.8
PBMC PWM	10.0	Lung fibroblast IL-4	5.1
PBMC PHA-L	5.8	Lung fibroblast IL-9	6.7

Ramos (B cell) none	18.3	Lung fibroblast IL-13	4.2
Ramos (B cell) ionomycin	17.8	Lung fibroblast IFN gamma	6.3
B lymphocytes PWM	10.7	Dermal fibroblast CCD1070 rest	7.0
B lymphocytes CD40L and IL-4	10.4	Dermal fibroblast CCD1070 TNF alpha	12.9
EOL-1 dbcAMP	5.7	Dermal fibroblast CCD1070 IL-1 beta	8.4
EOL-1 dbcAMP PMA/ionomycin	4.8	Dermal fibroblast IFN gamma	3.8
Dendritic cells none	8.0	Dermal fibroblast IL-4	7.6
Dendritic cells LPS	6.6	Dermal Fibroblasts rest	5.2
Dendritic cells anti-CD40	11.1	Neutrophils TNFa+LPS	4.1
Monocytes rest	17.1	Neutrophils rest	12.2
Monocytes LPS	9.1	Colon	1.6
Macrophages rest	6.9	Lung	3.5
Macrophages LPS	4.2	Thymus	51.1
HUVEC none	3.9	Kidney	14.8
HUVEC starved	5.5		

Table AIE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4043, Run 268362932	Tissue Name	Rel. Exp.(%) Ag4043, Run 268362932
Colon cancer 1	22.2	Bladder cancer NAT 2	0.4
Colon NAT 1	5.1	Bladder cancer NAT 3	0.6
Colon cancer 2	29.7	Bladder cancer NAT 4	11.6
Colon cancer NAT 2	7.9	Adenocarcinoma of the prostate 1	20.0
Colon cancer 3	67.8	Adenocarcinoma of the prostate 2	2.8
Colon cancer NAT 3	17.6	Adenocarcinoma of the prostate 3	9.2
Colon malignant cancer 4	91.4	Adenocarcinoma of the prostate 4	95.9
Colon normal adjacent tissue 4	5.3	Prostate cancer NAT 5	3.3
Lung cancer 1	9.7	Adenocarcinoma of the prostate 6	4.2
Lung NAT 1	3.8	Adenocarcinoma of the prostate 7	3.8
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	0.6

Lung NAT 2	23.5	Adenocarcinoma of the prostate 9	19.1
Squamous cell carcinoma 3	42.3	Prostate cancer NAT 10	1.8
Lung NAT 3	1.1	Kidney cancer 1	17.1
metastatic melanoma 1	12.9	Kidney NAT 1	8.8
Melanoma 2	11.7	Kidney cancer 2	74.2
Melanoma 3	18.2	Kidney NAT 2	20.0
metastatic melanoma 4	47.6	Kidney cancer 3	43.2
metastatic melanoma 5	42.9	Kidney NAT 3	8.4
Bladder cancer 1	2.6	Kidney cancer 4	14.6
Bladder cancer NAT 1	0.0	Kidney NAT 4	10.9
Bladder cancer 2	12.2		

CNS_neurodegeneration_v1.0 Summary: Ag4043 This panel does not show differential expression of the CG96231-01 gene in Alzheimer's disease. However, this expression profile confirms the presence of this gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system.

General_screening_panel_v1.4 Summary: Ag4043 Expression of the CG96231-01 gene is highest in fetal liver (CT=25.6). This gene is also expressed at moderate levels in other metabolic tissues, including pancreas, thyroid, adrenal, pituitary, adipose, and adult and fetal heart and skeletal muscle. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

In addition, this gene is expressed at much higher levels in fetal liver when compared to expression in the adult counterpart (CT=33.1). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue. In addition, the relative overexpression of this gene in fetal liver suggests that the protein product may act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of liver disease.

This gene is also expressed at moderate levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful

in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Expression is also slightly higher in cell lines derived from gastric, breast, ovarian, and prostate cancers. Overall, this gene is ubiquitously expressed in this panel suggesting a
5 role for this gene product in cell growth and/or proliferation.

Panel 4.1D Summary: Ag4043 Expression of the CG96231-01 gene is highest in the KU-812 basophil cell line treated with PMA/ionomycin (CT=27.2), with ubiquitous expression in this panel. Basophils release histamines and other biological modifiers in response to allergens and play an important role in the pathology of asthma and
10 hypersensitivity reactions. Therefore, therapeutics designed against the putative protein encoded by this gene may reduce or inhibit inflammation by blocking basophil function in these diseases. In addition, these cells are a reasonable model for the inflammatory cells that take part in various inflammatory lung and bowel diseases, such as asthma, Crohn's disease, and ulcerative colitis. Therefore, therapeutics that modulate the function of this
15 gene product may reduce or eliminate the symptoms of patients suffering from asthma, Crohn's disease, and ulcerative colitis.

General oncology screening panel_v_2.4 Summary: Ag4043 Highest expression of the CG96231-01 gene on this panel is seen in lung cancer (CT=28.9). In addition, this gene is overexpressed in lung and colon cancer when compared to
20 expression in the normal adjacent tissue. Thus, expression of this gene could be used as a marker of these cancers. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of lung and colon cancer.

AJ. CG96260-01: Phospholipase A2

25 Expression of gene CG96260-01 was assessed using the primer-probe set Ag4048, described in Table AJA. Results of the RTQ-PCR runs are shown in Tables AJB, AJC and AJD.

Table AJA. Probe Name Ag4048

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-atgaaggctcattgccatcct-3'	20	85	338
Probe	TET-5'-cacccacagcagtttctggcagtt-3'-TAMRA	24	129	339

Reverse	5'-atgtgtttgacctcctctga-3'	21	153	340
---------	----------------------------	----	-----	-----

Table AJB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4048, Run 214292262	Tissue Name	Rel. Exp.(%) Ag4048, Run 214292262
AD 1 Hippo	0.0	Control (Path) 3 Temporal Ctx	30.6
AD 2 Hippo	30.1	Control (Path) 4 Temporal Ctx	61.1
AD 3 Hippo	0.0	AD 1 Occipital Ctx	37.1
AD 4 Hippo	0.0	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	0.0	AD 3 Occipital Ctx	0.0
AD 6 Hippo	47.6	AD 4 Occipital Ctx	0.0
Control 2 Hippo	0.0	AD 5 Occipital Ctx	0.0
Control 4 Hippo	34.4	AD 6 Occipital Ctx	0.0
Control (Path) 3 Hippo	50.0	Control 1 Occipital Ctx	40.3
AD 1 Temporal Ctx	0.0	Control 2 Occipital Ctx	24.8
AD 2 Temporal Ctx	59.5	Control 3 Occipital Ctx	0.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	0.0
AD 4 Temporal Ctx	0.0	Control (Path) 1 Occipital Ctx	58.2
AD 5 Inf Temporal Ctx	0.0	Control (Path) 2 Occipital Ctx	0.0
AD 5 Sup Temporal Ctx	0.0	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	0.0	Control (Path) 4 Occipital Ctx	0.0
AD 6 Sup Temporal Ctx	59.9	Control 1 Parietal Ctx	84.7
Control 1 Temporal Ctx	0.0	Control 2 Parietal Ctx	28.7
Control 2 Temporal Ctx	100.0	Control 3 Parietal Ctx	0.0
Control 3 Temporal Ctx	0.0	Control (Path) 1 Parietal Ctx	18.7
Control 3 Temporal Ctx	27.0	Control (Path) 2 Parietal Ctx	60.3
Control (Path) 1 Temporal Ctx	0.0	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	45.4	Control (Path) 4 Parietal Ctx	0.0

Table AJC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4048, Run 218535022	Tissue Name	Rel. Exp.(%) Ag4048, Run 218535022
Adipose	9.3	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0

Melanoma* SK-MEL-5	0.0	Colon ca. SW480	23.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	65.1	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	52.9	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	25.5
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	51.4
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	92.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	19.2
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	12.2
Ovary	9.9	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	35.6
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	55.9	Thymus Pool	32.3
Trachea	22.7	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	27.9	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	4.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	15.4
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	22.5
Lung ca. HOP-62	7.2	Cerebral Cortex Pool	0.0
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.0
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.0	Brain (whole)	0.0
Liver ca. HepG2	0.0	Spinal Cord Pool	48.3

Kidney Pool	62.9	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	100.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	11.3
Renal ca. UO-31	0.0	Pancreas Pool	49.3

Table AJD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4048, Run 268362939	Tissue Name	Rel. Exp.(%) Ag4048, Run 268362939
Colon cancer 1	0.0	Bladder cancer NAT 2	3.9
Colon cancer NAT 1	16.2	Bladder cancer NAT 3	0.0
Colon cancer 2	41.8	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	39.8	Adenocarcinoma of the prostate 1	76.3
Colon cancer 3	4.6	Adenocarcinoma of the prostate 2	20.7
Colon cancer NAT 3	20.6	Adenocarcinoma of the prostate 3	0.0
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	5.8	Prostate cancer NAT 5	5.5
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	5.4
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	19.9
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	100.0
Squamous cell carcinoma 3	0.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	1.9
metastatic melanoma 1	0.0	Kidney NAT 1	0.0
Melanoma 2	0.0	Kidney cancer 2	21.6
Melanoma 3	7.7	Kidney NAT 2	0.0
metastatic melanoma 4	6.0	Kidney cancer 3	0.0
metastatic melanoma 5	5.6	Kidney NAT 3	0.0
Bladder cancer 1	4.4	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4048 This panel confirms the expression of the CG96260-01 gene at low levels in the temporal cortex of brain (CT=34.3). However, no differential expression of this gene was detected between

5 Alzheimer's diseased postmortem brains and those of non-demented controls in this

experiment. The the CG96260-01 gene codes for a homolog of rat phospholipase A2 (PLA(2)). Mouse deficient in cytosolic PLA(2), a member of PLA(2) family, has been shown to suffer smaller infarcts and fewer neurological deficits after transient occlusion of the middle cerebral artery and have less injury after administration of a dopaminergic selective neurotoxin (Sapirstein A, and Bonventre JV, 2000, Biochim Biophys Acta 2000 Oct 31;1488(1-2):139-48, PMID: 11080683). Thus, PLA(2) encoded by this gene could also play a role in inflammation and injuries of brain and pharmacological targeting of this enzyme may have important therapeutic benefits.

General_screening_panel_v1.4 Summary: Ag4048 Highest expression of the CG96260-01 gene is detected in renal cancer A498 cell line (CT=33). Therefore, therapeutic modulation of this gene product may be beneficial in the treatment of renal cancer.

In addition, moderate to low expression of this gene is seen in samples derived from normal tissues including testis, prostate, breast, kidney, lymphnode, thymus and spinal cord. Therefore, therapeutic modulation of this gene product may be useful in treatment of diseases associated with these tissues.

Among tissues with metabolic or endocrine function, this gene is expressed at moderate to low levels in pancreas, and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

The the CG96260-01 gene codes for a homolog of rat phospholipase A2 (PLA(2)). PLA(2) has been implicated in the pathogenesis and pathophysiology of acute pancreatitis (Nevalainen et al., 1999, Hepatogastroenterology 46(29):2731-5, PMID: 10576338). In addition, in blood plasma, PLA(2) modifies the circulating lipoproteins and so induce formation of small dense LDL particles, which are associated with increased risk for cardiovascular disease (Hurt-Camejo et al, 2001, Circ Res 89(4):298-304, PMID: 11509445). Furthermore, the cytoplasmic PLA(2) knockout mouse has revealed important roles of cPLA(2) in normal fertility, generation of eicosanoids from inflammatory cells, brain injuries and allergic responses (Sapirstein A, and Bonventre JV, 2000, Biochim Biophys Acta 2000 Oct 31;1488(1-2):139-48, PMID: 11080683). Therefore, therapeutic modulation of PLA(2) encoded by this gene may be useful in the treatment of pancreatitis, neurological disorders, allergies and cardiovascular diseases.

Panel 4.1D Summary: Ag4048 Expression of the CG96260-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel CNS_1 Summary: Ag4048 Expression of the CG96260-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

5 **General oncology screening panel_v_2.4 Summary:** Ag4048 Highest expression of the CG96260-01 gene is detected in prostate adenocarcinoma samples (CTs=33). In addition, low expression of this gene is also seen in a single colon cancer and the adjacent normal tissue sample. The CG96260-01 gene codes for a homolog of rat phospholipase A2 (PLA(2)). Phospholipase A2 has recently been recognized to be
10 involved in a wide number of pathophysiological situations, ranging from systemic and acute inflammatory conditions to cancer (Balsinde et al., 1999, Annu Rev Pharmacol Toxicol 39:175-89, PMID: 10331081). Therefore, therapeutic modulation of this gene product through the use of small molecule drug may be useful in the treatment of colon and prostate cancers. [mpattu, 29-Mar-02]

15

AK. CG96364-01: ADP/ATP Translocase 2 Like protein

Expression of gene CG96364-01 was assessed using the primer-probe set Ag4073, described in Table AKA. Results of the RTQ-PCR runs are shown in Tables AKB, AKC, AKD, AKE and AKF.

20

Table AKA. Probe Name Ag4073

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tcctcagatacttccccacc-3'	20	263	341
Probe	TET-5'-aacacaagcagatcttctctgggtggt-3'-TAMRA	26	311	342
Reverse	5'-aactgggtcctcttgtccac-3'	20	337	343

Table AKB. AI_comprehensive panel_v1.0

Tissue Name	Rel. Exp.(%) Ag4073, Run 248081917	Tissue Name	Rel. Exp.(%) Ag4073, Run 248081917
110967 COPD-F	11.3	112427 Match Control Psoriasis-F	40.6
110980 COPD-F	8.1	112418 Psoriasis-M	17.6
110968 COPD-M	12.1	112723 Match Control Psoriasis-M	8.8
110977 COPD-M	20.9	112419 Psoriasis-M	24.1

110989 Emphysema-F	39.0	112424 Match Control Psoriasis-M	14.7
110992 Emphysema-F	16.4	112420 Psoriasis-M	34.9
110993 Emphysema-F	12.3	112425 Match Control Psoriasis-M	27.0
110994 Emphysema-F	5.9	104689 (MF) OA Bone-Backus	31.0
110995 Emphysema-F	39.5	104690 (MF) Adj "Normal" Bone-Backus	8.4
110996 Emphysema-F	8.5	104691 (MF) OA Synovium-Backus	26.2
110997 Asthma-M	7.6	104692 (BA) OA Cartilage-Backus	18.9
111001 Asthma-F	18.7	104694 (BA) OA Bone-Backus	38.7
111002 Asthma-F	24.7	104695 (BA) Adj "Normal" Bone-Backus	17.6
111003 Atopic Asthma-F	29.1	104696 (BA) OA Synovium-Backus	32.3
111004 Atopic Asthma-F	40.6	104700 (SS) OA Bone-Backus	18.2
111005 Atopic Asthma-F	20.6	104701 (SS) Adj "Normal" Bone-Backus	19.3
111006 Atopic Asthma-F	6.2	104702 (SS) OA Synovium-Backus	34.2
111417 Allergy-M	13.0	117093 OA Cartilage Rep7	11.6
112347 Allergy-M	0.2	112672 OA Bone5	27.4
112349 Normal Lung-F	0.2	112673 OA Synovium5	12.3
112357 Normal Lung-F	26.6	112674 OA Synovial Fluid cells5	12.6
112354 Normal Lung-M	12.4	117100 OA Cartilage Rep14	8.1
112374 Crohns-F	19.3	112756 OA Bone9	100.0
112389 Match Control Crohns-F	21.6	112757 OA Synovium9	3.8
112375 Crohns-F	15.5	112758 OA Synovial Fluid Cells9	11.8
112732 Match Control Crohns-F	42.6	117125 RA Cartilage Rep2	15.3
112725 Crohns-M	7.5	113492 Bone2 RA	13.2
112387 Match Control Crohns-M	13.3	113493 Synovium2 RA	6.0
112378 Crohns-M	0.1	113494 Syn Fluid Cells RA	12.2
112390 Match Control Crohns-M	23.3	113499 Cartilage4 RA	23.5
112726 Crohns-M	24.0	113500 Bone4 RA	22.4
112731 Match Control Crohns-M	11.8	113501 Synovium4 RA	17.2
112380 Ulcer Col-F	14.4	113502 Syn Fluid Cells4 RA	10.4
112734 Match Control Ulcer Col-F	71.7	113495 Cartilage3 RA	13.4
112384 Ulcer Col-F	33.7	113496 Bone3 RA	13.3

112737 Match Control Ulcer Col-F	9.9	113497 Synovium3 RA	7.1
112386 Ulcer Col-F	1.6	113498 Syn Fluid Cells3 RA	16.0
112738 Match Control Ulcer Col-F	38.4	117106 Normal Cartilage Rep20	5.3
112381 Ulcer Col-M	0.3	113663 Bone3 Normal	0.6
112735 Match Control Ulcer Col-M	9.3	113664 Synovium3 Normal	0.0
112382 Ulcer Col-M	26.4	113665 Syn Fluid Cells3 Normal	0.3
112394 Match Control Ulcer Col-M	3.6	117107 Normal Cartilage Rep22	3.0
112383 Ulcer Col-M	19.5	113667 Bone4 Normal	6.1
112736 Match Control Ulcer Col-M	14.2	113668 Synovium4 Normal	10.2
112423 Psoriasis-F	19.3	113669 Syn Fluid Cells4 Normal	13.6

Table AKC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4073, Run 214294826	Tissue Name	Rel. Exp.(%) Ag4073, Run 214294826
AD 1 Hippo	8.4	Control (Path) 3 Temporal Ctx	3.8
AD 2 Hippo	18.3	Control (Path) 4 Temporal Ctx	14.2
AD 3 Hippo	2.1	AD 1 Occipital Ctx	8.9
AD 4 Hippo	3.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	100.0	AD 3 Occipital Ctx	3.9
AD 6 Hippo	28.1	AD 4 Occipital Ctx	12.4
Control 2 Hippo	12.9	AD 5 Occipital Ctx	25.2
Control 4 Hippo	6.5	AD 6 Occipital Ctx	77.4
Control (Path) 3 Hippo	3.2	Control 1 Occipital Ctx	3.3
AD 1 Temporal Ctx	3.5	Control 2 Occipital Ctx	58.2
AD 2 Temporal Ctx	17.4	Control 3 Occipital Ctx	12.8
AD 3 Temporal Ctx	2.3	Control 4 Occipital Ctx	3.9
AD 4 Temporal Ctx	11.1	Control (Path) 1 Occipital Ctx	49.3
AD 5 Inf Temporal Ctx	73.7	Control (Path) 2 Occipital Ctx	6.5
AD 5 SupTemporal Ctx	37.9	Control (Path) 3 Occipital Ctx	1.6
AD 6 Inf Temporal Ctx	27.0	Control (Path) 4 Occipital Ctx	8.6
AD 6 Sup Temporal Ctx	28.9	Control 1 Parietal Ctx	3.4
Control 1 Temporal Ctx	4.1	Control 2 Parietal Ctx	26.8
Control 2 Temporal Ctx	29.5	Control 3 Parietal Ctx	14.5
Control 3 Temporal Ctx	8.4	Control (Path) 1 Parietal Ctx	66.0

Control 4 Temporal Ctx	6.3	Control (Path) 2 Parietal Ctx	15.7
Control (Path) 1 Temporal Ctx	26.6	Control (Path) 3 Parietal Ctx	3.7
Control (Path) 2 Temporal Ctx	26.4	Control (Path) 4 Parietal Ctx	26.8

Table AKD. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4073, Run 218906344	Tissue Name	Rel. Exp.(%) Ag4073, Run 218906344
Adipose	1.4	Renal ca. TK-10	21.0
Melanoma* Hs688(A).T	5.9	Bladder	11.8
Melanoma* Hs688(B).T	6.7	Gastric ca. (liver met.) NCI-N87	22.5
Melanoma* M14	72.7	Gastric ca. KATO III	100.0
Melanoma* LOXIMVI	34.6	Colon ca. SW-948	21.6
Melanoma* SK-MEL-5	67.8	Colon ca. SW480	36.9
Squamous cell carcinoma SCC-4	22.7	Colon ca.* (SW480 met) SW620	66.0
Testis Pool	2.0	Colon ca. HT29	34.6
Prostate ca.* (bone met) PC-3	40.9	Colon ca. HCT-116	68.3
Prostate Pool	1.2	Colon ca. CaCo-2	24.3
Placenta	3.1	Colon cancer tissue	35.1
Uterus Pool	0.4	Colon ca. SW1116	7.5
Ovarian ca. OVCAR-3	36.6	Colon ca. Colo-205	23.8
Ovarian ca. SK-OV-3	32.8	Colon ca. SW-48	25.9
Ovarian ca. OVCAR-4	37.9	Colon Pool	3.2
Ovarian ca. OVCAR-5	24.1	Small Intestine Pool	1.7
Ovarian ca. IGROV-1	18.9	Stomach Pool	2.7
Ovarian ca. OVCAR-8	14.2	Bone Marrow Pool	0.8
Ovary	2.5	Fetal Heart	4.2
Breast ca. MCF-7	39.2	Heart Pool	1.5
Breast ca. MDA-MB-231	39.2	Lymph Node Pool	3.4
Breast ca. BT 549	52.1	Fetal Skeletal Muscle	0.9
Breast ca. T47D	39.8	Skeletal Muscle Pool	1.2
Breast ca. MDA-N	52.9	Spleen Pool	2.4
Breast Pool	2.7	Thymus Pool	3.6
Trachea	4.1	CNS cancer (glio/astro) U87-MG	32.1
Lung	0.8	CNS cancer (glio/astro) U-118-MG	27.2

Fetal Lung	4.9	CNS cancer (neuro;met) SK-N-AS	25.7
Lung ca. NCI-N417	17.4	CNS cancer (astro) SF-539	31.4
Lung ca. LX-1	48.3	CNS cancer (astro) SNB-75	50.3
Lung ca. NCI-H146	11.6	CNS cancer (glio) SNB-19	20.7
Lung ca. SHP-77	30.6	CNS cancer (glio) SF-295	16.5
Lung ca. A549	52.1	Brain (Amygdala) Pool	3.8
Lung ca. NCI-H526	12.4	Brain (cerebellum)	6.7
Lung ca. NCI-H23	31.6	Brain (fetal)	3.2
Lung ca. NCI-H460	15.3	Brain (Hippocampus) Pool	4.2
Lung ca. HOP-62	11.8	Cerebral Cortex Pool	5.1
Lung ca. NCI-H522	17.7	Brain (Substantia nigra) Pool	6.6
Liver	2.0	Brain (Thalamus) Pool	6.3
Fetal Liver	8.0	Brain (whole)	4.1
Liver ca. HepG2	21.8	Spinal Cord Pool	2.6
Kidney Pool	3.0	Adrenal Gland	4.6
Fetal Kidney	3.3	Pituitary gland Pool	0.8
Renal ca. 786-0	15.2	Salivary Gland	4.4
Renal ca. A498	5.7	Thyroid (female)	2.1
Renal ca. ACHN	18.9	Pancreatic ca. CAPAN2	30.4
Renal ca. UO-31	12.9	Pancreas Pool	4.5

Table AKE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4073, Run 171808865	Tissue Name	Rel. Exp.(%) Ag4073, Run 171808865
Secondary Th1 act	88.3	HUVEC IL-1beta	33.7
Secondary Th2 act	82.9	HUVEC IFN gamma	44.1
Secondary Tr1 act	59.5	HUVEC TNF alpha + IFN gamma	24.3
Secondary Th1 rest	12.8	HUVEC TNF alpha + IL4	32.1
Secondary Th2 rest	16.6	HUVEC IL-11	17.1
Secondary Tr1 rest	26.2	Lung Microvascular EC none	50.7
Primary Th1 act	86.5	Lung Microvascular EC TNFalpha + IL-1beta	31.6
Primary Th2 act	95.3	Microvascular Dermal EC none	29.3
Primary Tr1 act	100.0	Microvascular Dermal EC TNFalpha + IL-1beta	28.1
Primary Th1 rest	25.7	Bronchial epithelium TNFalpha + IL1beta	44.4

Primary Th2 rest	13.9	Small airway epithelium none	20.3
Primary Tr1 rest	34.9	Small airway epithelium TNFalpha + IL-1 beta	51.4
CD45RA CD4 lymphocyte act	77.4	Coronary artery SMC rest	17.7
CD45RO CD4 lymphocyte act	94.6	Coronary artery SMC TNFalpha + IL-1 beta	21.3
CD8 lymphocyte act	90.1	Astrocytes rest	14.5
Secondary CD8 lymphocyte rest	80.1	Astrocytes TNFalpha + IL-1 beta	12.2
Secondary CD8 lymphocyte act	45.4	KU-812 (Basophil) rest	24.5
CD4 lymphocyte none	7.2	KU-812 (Basophil) PMA/ionomycin	59.5
2ry Th1/Th2/Tr1_anti-CD95 CH11	42.3	CCD1106 (Keratinocytes) none	64.6
LAK cells rest	57.0	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	46.0
LAK cells IL-2	42.9	Liver cirrhosis	6.8
LAK cells IL-2+IL-12	31.0	NCI-H292 none	61.6
LAK cells IL-2+IFN gamma	23.2	NCI-H292 IL-4	79.0
LAK cells IL-2+ IL-18	28.7	NCI-H292 IL-9	95.3
LAK cells PMA/ionomycin	55.9	NCI-H292 IL-13	79.6
NK Cells IL-2 rest	59.0	NCI-H292 IFN gamma	65.5
Two Way MLR 3 day	27.5	HPAEC none	24.7
Two Way MLR 5 day	57.4	HPAEC TNF alpha + IL-1 beta	44.4
Two Way MLR 7 day	38.2	Lung fibroblast none	16.7
PBMC rest	14.2	Lung fibroblast TNF alpha + IL-1 beta	14.2
PBMC PWM	55.9	Lung fibroblast IL-4	23.2
PBMC PHA-L	72.2	Lung fibroblast IL-9	35.4
Ramos (B cell) none	65.5	Lung fibroblast IL-13	23.5
Ramos (B cell) ionomycin	100.0	Lung fibroblast IFN gamma	28.3
B lymphocytes PWM	82.4	Dermal fibroblast CCD1070 rest	36.1
B lymphocytes CD40L and IL-4	57.8	Dermal fibroblast CCD1070 TNF alpha	65.1
EOL-1 dbcAMP	51.4	Dermal fibroblast CCD1070 IL-1 beta	37.6
EOL-1 dbcAMP PMA/ionomycin	33.2	Dermal fibroblast IFN gamma	26.8

Dendritic cells none	46.0	Dermal fibroblast IL-4	31.2
Dendritic cells LPS	31.6	Dermal Fibroblasts rest	19.6
Dendritic cells anti-CD40	48.0	Neutrophils TNFa+LPS	3.1
Monocytes rest	25.3	Neutrophils rest	4.2
Monocytes LPS	18.0	Colon	35.4
Macrophages rest	70.2	Lung	17.6
Macrophages LPS	14.7	Thymus	32.8
HUVEC none	39.2	Kidney	47.6
HUVEC starved	43.8		

Table AKF. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4073, Run 268389975	Tissue Name	Rel. Exp.(%) Ag4073, Run 268389975
Colon cancer 1	50.3	Bladder cancer NAT 2	0.1
Colon cancer NAT 1	17.7	Bladder cancer NAT 3	0.1
Colon cancer 2	17.8	Bladder cancer NAT 4	1.7
Colon cancer NAT 2	22.4	Adenocarcinoma of the prostate 1	6.9
Colon cancer 3	100.0	Adenocarcinoma of the prostate 2	0.3
Colon cancer NAT 3	23.8	Adenocarcinoma of the prostate 3	2.5
Colon malignant cancer 4	47.6	Adenocarcinoma of the prostate 4	7.8
Colon normal adjacent tissue 4	14.5	Prostate cancer NAT 5	1.4
Lung cancer 1	8.2	Adenocarcinoma of the prostate 6	1.1
Lung NAT 1	0.5	Adenocarcinoma of the prostate 7	1.3
Lung cancer 2	14.8	Adenocarcinoma of the prostate 8	0.4
Lung NAT 2	0.7	Adenocarcinoma of the prostate 9	3.7
Squamous cell carcinoma 3	23.0	Prostate cancer NAT 10	0.1
Lung NAT 3	0.4	Kidney cancer 1	6.3
metastatic melanoma 1	3.8	Kidney NAT 1	3.3
Melanoma 2	3.3	Kidney cancer 2	21.2
Melanoma 3	2.3	Kidney NAT 2	17.7
metastatic melanoma 4	8.8	Kidney cancer 3	4.5
metastatic melanoma 5	14.7	Kidney NAT 3	6.7
Bladder cancer 1	0.6	Kidney cancer 4	11.6
Bladder cancer NAT 1	0.0	Kidney NAT 4	17.4
Bladder cancer 2	1.9		

AI_comprehensive_panel_v1.0 Summary: Ag4073 Highest expression of the CG96364-01 gene is detected in osteoarthritic bone sample (CT=27). Low to moderate levels of expression of this gene are detected in samples derived from osteoarthritic (OA) bone and adjacent bone as well as OA cartilage, OA synovium and OA synovial fluid samples. Moderate level expression is also detected in cartilage, bone, synovium and synovial fluid samples from rheumatoid arthritis patients. Low level expression of this gene is also detected in normal samples derived from cartilage, synovium, bone or synovial fluid cells, normal lung samples, COPD lung, emphysema, atopic asthma, asthma, allergy, Crohn's disease (normal matched control and diseased), ulcerative colitis(normal matched control and diseased), and psoriasis (normal matched control and diseased). Therefore, therapeutic modulation of this gene product may ameliorate symptoms/conditions associated with autoimmune and inflammatory disorders including psoriasis, allergy, asthma, inflammatory bowel disease, rheumatoid arthritis and osteoarthritis

CNS_neurodegeneration_v1.0 Summary: Ag4073 This panel confirms the expression of the CG96364-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4073 Highest expression of the CG96364-01 gene is detected in a gastric cancer KATO III cell line (CT=22). High expression of this gene is seen in cluster of breast, ovarian, colon, gastric, renal, lung, pancreatic, CNS, hepatic, prostate cancer cell lines and melanoma cell lines. Thus, therapeutic modulation of this gene product could be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

The CG100466-01 gene codes for adenine nucleotide translocator (ANT 2) homolog. Dysfunctioning of the ANT2 have been shown to induce myopathies in mouse and in humans (Fiore et al., 2001, Clin Chim Acta 311(2):125-35, PMID: 11566172). ANT has a role in mtDNA maintenance and mutation in ANT has been
5 implicated in autosomal dominant progressive external ophthalmoplegia and other mitochondrial diseases (Kaukonen et al., 2000, Science 289(5480):782-5, PMID: 10926541). Mice deficient in the heart/muscle specific isoform of the ANT1 exhibit many of the hallmarks of human oxidative phosphorylation (OXPHOS) disease, including a dramatic proliferation of skeletal muscle mitochondria (Murdoch et al., 1999, J Biol Chem
10 274(20):14429-33, PMID: 10318868). Therefore, therapeutic modulation of the ANT protein encoded by the CG100466-01 gene through the use of small molecule drug could be useful in the treatment mitochondria related diseases.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus,
15 cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4073 Highest expression of the CG96364-01 gene is detected in activated primary Tr1 and ionomycin treated Ramos B cell (CTs=25.4). This
20 gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of
25 expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead
30 to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4073 Highest expression of the CG96364-01 gene is detected in colon cancer OD06297-04 sample (CT=24.3). Higher expression of this gene is seen in colon cancer and lung cancer compared to the normal adjacent tissues. Therapeutic modulation of this gene product could be beneficial in the treatment of these cancers. Please see Panel 1.4 for additional discussion of the potential utility of this gene in treatment of cancers.

AL. CG96422-01: ADP/ATP Translocase 3

Expression of gene CG96422-01 was assessed using the primer-probe set Ag4057, described in Table ALA. Results of the RTQ-PCR runs are shown in Tables ALB and ALC.

Table ALA. Probe Name Ag4057

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gccaatgtcatccgctact-3'	19	232	344
Probe	TET-5'-agccctcaatttcgccttcaaggata-3'-TAMRA	26	261	345
Reverse	5'-gccaggaagatctgcttgctact-3'	22	287	346

Table ALB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4057, Run 171620020	Tissue Name	Rel. Exp.(%) Ag4057, Run 171620020
Secondary Th1 act	33.9	HUVEC IL-1beta	51.1
Secondary Th2 act	32.5	HUVEC IFN gamma	56.6
Secondary Tr1 act	29.9	HUVEC TNF alpha + IFN gamma	29.9
Secondary Th1 rest	14.9	HUVEC TNF alpha + IL4	42.9
Secondary Th2 rest	21.5	HUVEC IL-11	43.5
Secondary Tr1 rest	22.7	Lung Microvascular EC none	81.8
Primary Th1 act	33.7	Lung Microvascular EC TNFalpha + IL-1beta	60.3
Primary Th2 act	54.7	Microvascular Dermal EC none	53.6
Primary Tr1 act	42.6	Microvascular Dermal EC TNFalpha + IL-1beta	36.3
Primary Th1 rest	16.4	Bronchial epithelium TNFalpha + IL1beta	65.1
Primary Th2 rest	8.6	Small airway epithelium none	21.8

Primary Tr1 rest	22.2	Small airway epithelium TNFalpha + IL-1beta	36.9
CD45RA CD4 lymphocyte act	33.0	Coronary artery SMC rest	46.7
CD45RO CD4 lymphocyte act	47.0	Coronary artery SMC TNFalpha + IL-1beta	39.5
CD8 lymphocyte act	55.1	Astrocytes rest	29.3
Secondary CD8 lymphocyte rest	76.3	Astrocytes TNFalpha + IL-1beta	26.8
Secondary CD8 lymphocyte act	25.5	KU-812 (Basophil) rest	72.7
CD4 lymphocyte none	26.2	KU-812 (Basophil) PMA/ionomycin	100.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	34.6	CCD1106 (Keratinocytes) none	45.4
LAK cells rest	26.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	27.4
LAK cells IL-2	0.0	Liver cirrhosis	8.1
LAK cells IL-2+IL-12	29.5	NCI-H292 none	90.1
LAK cells IL-2+IFN gamma	22.1	NCI-H292 IL-4	81.8
LAK cells IL-2+ IL-18	17.7	NCI-H292 IL-9	92.7
LAK cells PMA/ionomycin	41.5	NCI-H292 IL-13	63.3
NK Cells IL-2 rest	28.7	NCI-H292 IFN gamma	57.4
Two Way MLR 3 day	24.0	HPAEC none	50.0
Two Way MLR 5 day	33.0	HPAEC TNF alpha + IL-1 beta	65.5
Two Way MLR 7 day	28.1	Lung fibroblast none	49.0
PBMC rest	32.1	Lung fibroblast TNF alpha + IL-1 beta	14.2
PBMC PWM	33.0	Lung fibroblast IL-4	37.6
PBMC PHA-L	30.1	Lung fibroblast IL-9	57.4
Ramos (B cell) none	36.1	Lung fibroblast IL-13	84.7
Ramos (B cell) ionomycin	51.8	Lung fibroblast IFN gamma	51.1
B lymphocytes PWM	42.9	Dermal fibroblast CCD1070 rest	46.7
B lymphocytes CD40L and IL-4	52.1	Dermal fibroblast CCD1070 TNF alpha	53.2
EOL-1 dbcAMP	46.7	Dermal fibroblast CCD1070 IL-1 beta	34.4
EOL-1 dbcAMP PMA/ionomycin	53.2	Dermal fibroblast IFN gamma	25.5
Dendritic cells none	57.0	Dermal fibroblast IL-4	55.5

Dendritic cells LPS	38.2	Dermal Fibroblasts rest	21.2
Dendritic cells anti-CD40	66.9	Neutrophils TNFa+LPS	4.2
Monocytes rest	51.4	Neutrophils rest	6.5
Monocytes LPS	16.7	Colon	14.8
Macrophages rest	45.7	Lung	23.7
Macrophages LPS	12.7	Thymus	40.3
HUVEC none	38.7	Kidney	36.3
HUVEC starved	57.0		

Table ALC. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4057, Run 268390123	Tissue Name	Rel. Exp.(%) Ag4057, Run 268390123
Colon cancer 1	9.2	Bladder cancer NAT 2	0.0
Colon NAT 1	6.5	Bladder cancer NAT 3	0.1
Colon cancer 2	12.8	Bladder cancer NAT 4	3.4
Colon cancer NAT 2	12.7	Adenocarcinoma of the prostate 1	11.6
Colon cancer 3	43.8	Adenocarcinoma of the prostate 2	0.9
Colon cancer NAT 3	16.2	Adenocarcinoma of the prostate 3	4.4
Colon malignant cancer 4	100.0	Adenocarcinoma of the prostate 4	14.4
Colon normal adjacent tissue 4	8.7	Prostate cancer NAT 5	1.9
Lung cancer 1	8.0	Adenocarcinoma of the prostate 6	3.3
Lung NAT 1	0.4	Adenocarcinoma of the prostate 7	4.0
Lung cancer 2	19.1	Adenocarcinoma of the prostate 8	0.6
Lung NAT 2	0.6	Adenocarcinoma of the prostate 9	17.6
Squamous cell carcinoma 3	39.0	Prostate cancer NAT 10	0.3
Lung NAT 3	0.3	Kidney cancer 1	6.5
metastatic melanoma 1	16.4	Kidney NAT 1	3.3
Melanoma 2	2.7	Kidney cancer 2	42.3
Melanoma 3	1.9	Kidney NAT 2	11.3
metastatic melanoma 4	44.1	Kidney cancer 3	13.1
metastatic melanoma 5	38.7	Kidney NAT 3	4.7
Bladder cancer 1	0.2	Kidney cancer 4	16.8
Bladder cancer NAT 1	0.0	Kidney NAT 4	15.4
Bladder cancer 2	2.5		

CNS_neurodegeneration_v1.0 Summary: Ag4057 Results from one experiment with the CG96422-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4057 Results from one
5 experiment with the CG96422-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 4.1D Summary: Ag4057 Highest expression of the CG96422-01 gene is detected in PMA/ionomycin treated basophils (CT27.4). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in
10 health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and
15 tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory
20 diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4057 Highest expression of the CG96422-01 gene is detected in colon malignant cancer sample (CT=27.5). In addition, significant expression of this gene is associated with kidney
25 cancer, prostate adenocarcinoma, bladder cancer, melanoma, lung and colon cancers. Therefore, therapeutic modulation of this gene product through the use of small molecule target may be beneficial in the treatment of these cancers.

AM. CG96442-01: Acyl CoA-Domain Protein

30 Expression of gene CG96442-01 was assessed using the primer-probe set Ag4058, described in Table AMA. Results of the RTQ-PCR runs are shown in Tables AMB and AMC.

Table AMA. Probe Name Ag4058

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - tcaaaggagattgtgccaatta - 3'	22	703	347
Probe	TET-5' - tgcacctgtcagtgttcaacaacaga - 3' - TAMRA	26	726	348
Reverse	5' - ctgtttgggtttccgaattaat - 3'	22	756	349

Table AMB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4058, Run 218902477	Tissue Name	Rel. Exp.(%) Ag4058, Run 218902477
Adipose	2.8	Renal ca. TK-10	9.9
Melanoma* Hs688(A).T	0.0	Bladder	33.2
Melanoma* Hs688(B).T	5.5	Gastric ca. (liver met.) NCI-N87	26.2
Melanoma* M14	0.0	Gastric ca. KATO.III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	17.3	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	1.6
Testis Pool	66.4	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	2.0
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	19.9
Uterus Pool	2.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	61.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	3.2	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	49.7	Colon Pool	16.0
Ovarian ca. OVCAR-5	1.8	Small Intestine Pool	5.2
Ovarian ca. IGROV-1	0.0	Stomach Pool	19.9
Ovarian ca. OVCAR-8	5.4	Bone Marrow Pool	6.7
Ovary	6.3	Fetal Heart	8.2
Breast ca. MCF-7	3.8	Heart Pool	2.4
Breast ca. MDA-MB-231	7.4	Lymph Node Pool	15.3
Breast ca. BT 549	4.0	Fetal Skeletal Muscle	10.4
Breast ca. T47D	11.3	Skeletal Muscle Pool	1.2
Breast ca. MDA-N	0.0	Spleen Pool	100.0
Breast Pool	28.3	Thymus Pool	5.8
Trachea	1.1	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0

Fetal Lung	13.5	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	3.4	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	12.7	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	1.9	CNS cancer (glio) SF-295	3.4
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.0
Lung ca. NCI-H23	14.7	Brain (fetal)	5.8
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	6.6
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	8.0
Lung ca. NCI-H522	3.2	Brain (Substantia nigra) Pool	11.7
Liver	3.0	Brain (Thalamus) Pool	9.0
Fetal Liver	4.2	Brain (whole)	0.0
Liver ca. HepG2	0.0	Spinal Cord Pool	0.0
Kidney Pool	20.0	Adrenal Gland	16.4
Fetal Kidney	23.8	Pituitary gland Pool	8.0
Renal ca. 786-0	0.0	Salivary Gland	8.3
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	27.9	Pancreatic ca. CAPAN2	7.4
Renal ca. UO-31	21.2	Pancreas Pool	8.7

Table AMC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4058, Run 171620034	Tissue Name	Rel. Exp.(%) Ag4058, Run 171620034
Secondary Th1 act	0.0	HUVEC IL-1beta	7.7
Secondary Th2 act	0.0	HUVEC IFN gamma	2.9
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	2.9
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	2.5
Secondary Th2 rest	0.0	HUVEC IL-11	5.8
Secondary Tr1 rest	3.1	Lung Microvascular EC none	7.5
Primary Th1 act	1.7	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	3.7
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium	0.0

		TNFalpha + IL-1beta	
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	4.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	9.5	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	1.8	Liver cirrhosis	9.3
LAK cells IL-2+IL-12	2.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	6.5	HPAEC none	5.2
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	48.3
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	2.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	3.5	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	4.0

Monocytes rest	38.4	Neutrophils rest	0.0
Monocytes LPS	52.5	Colon	0.0
Macrophages rest	9.9	Lung	0.0
Macrophages LPS	5.9	Thymus	6.1
HUVEC none	0.0	Kidney	100.0
HUVEC starved	2.5		

AI_comprehensive_panel_v1.0 Summary: Ag4058 Expression of the CG96442-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

CNS_neurodegeneration_v1.0 Summary: Ag4058 Results from one experiment with the CG96442-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4058 Highest expression of the CG96442-01 gene is detected in spleen (CT=31.4). In addition, moderate to low expression of this gene is seen in pancreas, adrenal gland, gastrointestinal tract, kidney, and testis. Therefore, therapeutic modulation of this gene can be useful in the treatment of diseases associated with these tissues, including diabetes, obesity, inflammatory bowel disease, lupus and glomerulonephritis.

Low expression of this gene is also detected in brain thalamus and substantia nigra samples. Therefore, therapeutic modulation of this gene product can be beneficial in neurological disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Low expression of this gene is associated with cluster of cancer cell lines including renal cancer, gastric cancer, breast cancer, ovarian cancer, lung cancer and melanoma cell lines and colon cancer tissue. Therefore, therapeutic modulation of this gene product can be beneficial in the treatment of these cancers.

Interestingly, this gene is expressed at much higher levels in fetal (CT=34.5) when compared to adult skeletal muscle and lung (CT>38). Thus, expression of this gene can be used to distinguish fetal from adult skeletal muscle and lung. In addition, the relative overexpression of this gene in fetal skeletal muscle suggests that the protein product may enhance growth or development of muscle and lung in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of muscle and lung related diseases.

Panel 4.1D Summary: Ag4058 Highest expression of the CG96442-01 gene is detected in kidney (CT=31.4). Therefore, expression of this gene can be used to distinguish this sample from other samples in the panel. In addition, small molecule therapies designed with the protein encoded for by this gene could modulate kidney function and be important in the treatment of inflammatory or autoimmune diseases that affect the kidney, including lupus and glomerulonephritis.

Low level expression of this gene is also detected in TNF alpha + IL-1 beta treated HPAEC, liver cirrhosis, and monocytes. Therefore, therapeutic modulation of this gene product can be useful in the treatment of inflammatory and autoimmune diseases such as asthma, IBD, and psoriasis and liver cirrhosis.

General oncology screening panel_v_2.4 Summary: Ag4058 Expression of the CG96442-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

15 AN. CG96501-01: Brain Mitochondrial Carrier Protein-1 Like protein

Expression of gene CG96501-01 was assessed using the primer-probe set Ag4059, described in Table ANA. Results of the RTQ-PCR runs are shown in Tables ANB, ANC and AND.

Table ANA. Probe Name Ag4059

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctcagtttcccagtcattgatg-3'	22	53	350
Probe	TET-5'-cctgagcagcagcaaagtaccactgt-3'-TAMRA	26	77	351
Reverse	5'-ggcgtcatatacaaagggttt-3'	21	131	352

20

Table ANB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4059, Run 218902504	Tissue Name	Rel. Exp.(%) Ag4059, Run 218902504
Adipose	0.1	Renal ca. TK-10	1.4
Melanoma* Hs688(A).T	0.6	Bladder	1.2
Melanoma* Hs688(B).T	0.3	Gastric ca. (liver met.) NCI-N87	4.3
Melanoma* M14	0.6	Gastric ca. KATO III	1.8
Melanoma* LOXIMVI	0.4	Colon ca. SW-948	0.4
Melanoma* SK-MEL-5	0.7	Colon ca. SW480	1.7
Squamous cell	0.3	Colon ca.* (SW480 met) SW620	2.6

carcinoma SCC-4			
Testis Pool	1.1	Colon ca. HT29	1.3
Prostate ca.* (bone met) PC-3	1.0	Colon ca. HCT-116	2.6
Prostate Pool	0.5	Colon ca. CaCo-2	2.8
Placenta	0.7	Colon cancer tissue	1.7
Uterus Pool	0.3	Colon ca. SW1116	0.9
Ovarian ca. OVCAR-3	0.6	Colon ca. Colo-205	0.3
Ovarian ca. SK-OV-3	3.6	Colon ca. SW-48	0.1
Ovarian ca. OVCAR-4	0.5	Colon Pool	3.3
Ovarian ca. OVCAR-5	2.1	Small Intestine Pool	4.0
Ovarian ca. IGROV-1	0.6	Stomach Pool	1.5
Ovarian ca. OVCAR-8	0.6	Bone Marrow Pool	1.3
Ovary	2.4	Fetal Heart	1.2
Breast ca. MCF-7	2.0	Heart Pool	1.2
Breast ca. MDA-MB-231	3.3	Lymph Node Pool	3.7
Breast ca. BT 549	5.4	Fetal Skeletal Muscle	1.7
Breast ca. T47D	2.7	Skeletal Muscle Pool	1.1
Breast ca. MDA-N	0.6	Spleen Pool	1.8
Breast Pool	3.9	Thymus Pool	2.5
Trachea	1.3	CNS cancer (glio/astro) U87-MG	3.4
Lung	0.7	CNS cancer (glio/astro) U-118-MG	2.7
Fetal Lung	3.8	CNS cancer (neuro;met) SK-N-AS	3.8
Lung ca. NCI-N417	0.4	CNS cancer (astro) SF-539	1.0
Lung ca. LX-1	2.1	CNS cancer (astro) SNB-75	3.5
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.7
Lung ca. SHP-77	1.1	CNS cancer (glio) SF-295	3.3
Lung ca. A549	2.6	Brain (Amygdala) Pool	0.6
Lung ca. NCI-H526	0.2	Brain (cerebellum)	0.9
Lung ca. NCI-H23	2.9	Brain (fetal)	1.2
Lung ca. NCI-H460	2.3	Brain (Hippocampus) Pool	1.2
Lung ca. HOP-62	2.3	Cerebral Cortex Pool	0.8
Lung ca. NCI-H522	1.5	Brain (Substantia nigra) Pool	3.5
Liver	0.1	Brain (Thalamus) Pool	1.2
Fetal Liver	2.2	Brain (whole)	1.2
Liver ca. HepG2	2.1	Spinal Cord Pool	1.3

Kidney Pool	3.1	Adrenal Gland	0.7
Fetal Kidney	3.9	Pituitary gland Pool	0.5
Renal ca. 786-0	2.1	Salivary Gland	0.4
Renal ca. A498	1.0	Thyroid (female)	0.0
Renal ca. ACHN	1.3	Pancreatic ca. CAPAN2	2.1
Renal ca. UO-31	2.0	Pancreas Pool	100.0

Table ANC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4059, Run 171620235	Tissue Name	Rel. Exp.(%) Ag4059, Run 171620235
Secondary Th1 act	14.4	HUVEC IL-1beta	15.7
Secondary Th2 act	39.2	HUVEC IFN gamma	26.1
Secondary Tr1 act	34.4	HUVEC TNF alpha + IFN gamma	16.8
Secondary Th1 rest	18.8	HUVEC TNF alpha + IL4	10.6
Secondary Th2 rest	6.1	HUVEC IL-11	10.7
Secondary Tr1 rest	36.6	Lung Microvascular EC none	19.5
Primary Th1 act	13.4	Lung Microvascular EC TNFalpha + IL-1beta	14.4
Primary Th2 act	31.2	Microvascular Dermal EC none	18.8
Primary Tr1 act	22.2	Microvascular Dermal EC TNFalpha + IL-1beta	12.2
Primary Th1 rest	6.4	Bronchial epithelium TNFalpha + IL1beta	4.6
Primary Th2 rest	16.5	Small airway epithelium none	3.0
Primary Tr1 rest	12.1	Small airway epithelium TNFalpha + IL-1beta	4.4
CD45RA CD4 lymphocyte act	21.9	Coronary artery SMC rest	2.8
CD45RO CD4 lymphocyte act	39.5	Coronary artery SMC TNFalpha + IL-1beta	4.5
CD8 lymphocyte act	40.3	Astrocytes rest	4.3
Secondary CD8 lymphocyte rest	23.2	Astrocytes TNFalpha + IL-1beta	10.5
Secondary CD8 lymphocyte act	7.2	KU-812 (Basophil) rest	33.9
CD4 lymphocyte none	27.7	KU-812 (Basophil) PMA/ionomycin	34.6
2ry Th1/Th2/Tr1_anti- CD95 CH11	45.7	CCD1106 (Keratinocytes) none	9.2
LAK cells rest	17.9	CCD1106 (Keratinocytes)	8.7

		TNFalpha + IL-1beta	
LAK cells IL-2	27.7	Liver cirrhosis	4.3
LAK cells IL-2+IL-12	20.2	NCI-H292 none	11.0
LAK cells IL-2+IFN gamma	12.0	NCI-H292 IL-4	11.5
LAK cells IL-2+ IL-18	29.3	NCI-H292 IL-9	24.3
LAK cells PMA/ionomycin	5.6	NCI-H292 IL-13	12.3
NK Cells IL-2 rest	37.6	NCI-H292 IFN gamma	13.8
Two Way MLR 3 day	52.1	HPAEC none	18.8
Two Way MLR 5 day	22.8	HPAEC TNF alpha + IL-1 beta	20.0
Two Way MLR 7 day	30.8	Lung fibroblast none	23.8
PBMC rest	4.6	Lung fibroblast TNF alpha + IL-1 beta	9.2
PBMC PWM	24.5	Lung fibroblast IL-4	10.5
PBMC PHA-L	29.1	Lung fibroblast IL-9	23.8
Ramos (B cell) none	41.8	Lung fibroblast IL-13	19.1
Ramos (B cell) ionomycin	37.9	Lung fibroblast IFN gamma	20.9
B lymphocytes PWM	16.4	Dermal fibroblast CCD1070 rest	29.1
B lymphocytes CD40L and IL-4	42.3	Dermal fibroblast CCD1070 TNF alpha	23.7
EOL-1 dbcAMP	40.3	Dermal fibroblast CCD1070 IL-1 beta	4.4
EOL-1 dbcAMP PMA/ionomycin	15.5	Dermal fibroblast IFN gamma	7.3
Dendritic cells none	11.4	Dermal fibroblast IL-4	20.3
Dendritic cells LPS	18.9	Dermal Fibroblasts rest	21.3
Dendritic cells anti-CD40	4.2	Neutrophils TNFa+LPS	2.8
Monocytes rest	6.7	Neutrophils rest	17.4
Monocytes LPS	21.9	Colon	3.1
Macrophages rest	16.2	Lung	6.5
Macrophages LPS	12.9	Thymus	57.0
HUVEC none	12.7	Kidney	100.0
HUVEC starved	33.9		

Table AND. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4059, Run 268390135	Tissue Name	Rel. Exp.(%) Ag4059, Run 268390135
-------------	--	-------------	--

Colon cancer 1	8.7	Bladder cancer NAT 2	0.0
Colon NAT 1	0.6	Bladder cancer NAT 3	0.4
Colon cancer 2	9.5	Bladder cancer NAT 4	8.1
Colon cancer NAT 2	4.1	Adenocarcinoma of the prostate 1	31.2
Colon cancer 3	12.9	Adenocarcinoma of the prostate 2	2.4
Colon cancer NAT 3	14.7	Adenocarcinoma of the prostate 3	12.2
Colon malignant cancer 4	29.7	Adenocarcinoma of the prostate 4	17.0
Colon normal adjacent tissue 4	2.1	Prostate cancer NAT 5	2.7
Lung cancer 1	7.9	Adenocarcinoma of the prostate 6	4.7
Lung NAT 1	1.9	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	23.8	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.6	Adenocarcinoma of the prostate 9	27.4
Squamous cell carcinoma 3	11.2	Prostate cancer NAT 10	0.5
Lung NAT 3	0.0	Kidney cancer 1	17.0
metastatic melanoma 1	19.5	Kidney NAT 1	8.3
Melanoma 2	1.9	Kidney cancer 2	47.6
Melanoma 3	1.3	Kidney NAT 2	20.7
metastatic melanoma 4	97.9	Kidney cancer 3	44.1
metastatic melanoma 5	100.0	Kidney NAT 3	6.4
Bladder cancer 1	0.5	Kidney cancer 4	10.4
Bladder cancer NAT 1	0.0	Kidney NAT 4	2.1
Bladder cancer 2	3.4		

CNS_neurodegeneration_v1.0 Summary: Ag4059 Results from one experiment with the CG96501-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

- 5 **General_screening_panel_v1.4 Summary:** Ag4059 Highest expression of the CG96501-01 gene is detected in pancreas (CT=24.6). In addition, moderate expression of this gene is seen among the tissues with metabolic or endocrine function such as pancreas, adipose, adrenal gland, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the
- 10 treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

The CG96501-01 gene codes for homolog of mouse mitochondrial uncoupling protein 5 (UCP 5). Proteins belonging to UCP family play important role in

thermoregulatory responses, including acute hypothermia usually followed by temperature recovery (Yu et al., 2000, Am J Physiol Endocrinol Metab 279(2):E433-46). In addition, mutation in UCP3, another member of UCP family, causes severe obesity and type II diabetes (Brown et al., 1999, Hum. Mutat. 13: 508 only). Therefore, therapeutic modulation of this gene product can be useful in treatment of disorders affecting thermoregulatory responses, obesity and type II diabetes.

Significant expression of this gene is also seen in cluster of cancer cell lines including pancreatic cancer, CNS cancer, colon and gastric cancer, renal cancer, lung cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma and melanoma cell lines. Therefore, therapeutic modulation of this gene product can be beneficial in the treatment of these cancers.

Also, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4059 Highest expression of the CG96501-01 gene is detected in kidney (CT=32). This gene is expressed at low to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4059 Highest expression of the CG96501-01 gene is detected in tow metastatic melanoma samples (CTs=31.7). Interestingly, significant expression of this gene is also seen in cluster of cancer samples including kidney, prostate adenocarcinoma, metastatic melanoma, lung and colon cancers. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of these cancers.

AO. CG96557-01: L-ALLO-Threonine Aldolase-Like Protein

Expression of gene CG96557-01 was assessed using the primer-probe set Ag4067, described in Table AOA.

Table AOA. Probe Name Ag4067

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tggagcactgtgactctgtgt-3'	21	614	353
Probe	TET-5'-ctttctgcctctccaagggcctg-3'-TAMRA	23	635	354
Reverse	5'-aggtctcttcaatgaagtcctt-3'	22	691	355

CNS_neurodegeneration_v1.0 Summary: Ag4067 Results from one experiment with the CG96557-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4067 Expression of the CG96557-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 2.2 Summary: Ag4067 Expression of the CG96557-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4067 Expression of the CG96557-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown). Results from one experiment (run 171808741) with the CG96557-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

AP. CG96581-01: RP42

Expression of gene CG96581-01 was assessed using the primer-probe set Ag4070, described in Table APA. Results of the RTQ-PCR runs are shown in Tables APB, APC, APD and APE.

Table APA. Probe Name Ag4070

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ccagtagtcacaggtggaaaac-3'	22	835	356
Probe	TET-5'-cagccttttctaggcagcaagttaagca-3'-TAMRA	28	858	357
Reverse	5'-caatctccttgaggatacaaa-3'	22	906	358

5

Table APB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4070, Run 214294588	Tissue Name	Rel. Exp.(%) Ag4070, Run 214294588
AD 1 Hippo	10.2	Control (Path) 3 Temporal Ctx	0.1
AD 2 Hippo	14.1	Control (Path) 4 Temporal Ctx	26.6
AD 3 Hippo	6.4	AD 1 Occipital Ctx	16.8
AD 4 Hippo	4.0	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	100.0	AD 3 Occipital Ctx	5.1
AD 6 Hippo	55.1	AD 4 Occipital Ctx	15.6
Control 2 Hippo	15.3	AD 5 Occipital Ctx	12.8
Control 4 Hippo	10.1	AD 6 Occipital Ctx	28.9
Control (Path) 3 Hippo	4.5	Control 1 Occipital Ctx	3.5
AD 1 Temporal Ctx	13.5	Control 2 Occipital Ctx	76.8
AD 2 Temporal Ctx	19.8	Control 3 Occipital Ctx	19.3
AD 3 Temporal Ctx	4.1	Control 4 Occipital Ctx	5.2
AD 4 Temporal Ctx	16.8	Control (Path) 1 Occipital Ctx	82.9
AD 5 Inf Temporal Ctx	50.0	Control (Path) 2 Occipital Ctx	11.3
AD 5 Sup Temporal Ctx	20.0	Control (Path) 3 Occipital Ctx	1.6
AD 6 Inf Temporal Ctx	66.4	Control (Path) 4 Occipital Ctx	12.6
AD 6 Sup Temporal Ctx	62.4	Control 1 Parietal Ctx	3.7
Control 1 Temporal Ctx	4.2	Control 2 Parietal Ctx	21.5
Control 2 Temporal Ctx	32.5	Control 3 Parietal Ctx	17.8
Control 3 Temporal Ctx	9.9	Control (Path) 1 Parietal Ctx	58.6
Control 4 Temporal Ctx	4.9	Control (Path) 2 Parietal Ctx	21.0
Control (Path) 1 Temporal Ctx	37.1	Control (Path) 3 Parietal Ctx	2.4
Control (Path) 2	32.1	Control (Path) 4 Parietal Ctx	66.9

Temporal Ctx			
--------------	--	--	--

Table APC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4070, Run 172382436	Tissue Name	Rel. Exp.(%) Ag4070, Run 172382436
Secondary Th1 act	42.9	HUVEC IL-1beta	27.0
Secondary Th2 act	58.6	HUVEC IFN gamma	27.0
Secondary Tr1 act	47.3	HUVEC TNF alpha + IFN gamma	10.8
Secondary Th1 rest	16.2	HUVEC TNF alpha + IL4	21.3
Secondary Th2 rest	14.6	HUVEC IL-11	15.7
Secondary Tr1 rest	27.7	Lung Microvascular EC none	38.2
Primary Th1 act	46.7	Lung Microvascular EC TNFalpha + IL-1beta	31.0
Primary Th2 act	46.7	Microvascular Dermal EC none	26.1
Primary Tr1 act	42.9	Microvascular Dermal EC TNFalpha + IL-1beta	19.8
Primary Th1 rest	17.9	Bronchial epithelium TNFalpha + IL1beta	33.7
Primary Th2 rest	5.6	Small airway epithelium none	11.1
Primary Tr1 rest	19.6	Small airway epithelium TNFalpha + IL-1beta	22.4
CD45RA CD4 lymphocyte act	42.3	Coronary artery SMC rest	17.8
CD45RO CD4 lymphocyte act	48.0	Coronary artery SMC TNFalpha + IL-1beta	19.5
CD8 lymphocyte act	57.4	Astrocytes rest	14.1
Secondary CD8 lymphocyte rest	26.6	Astrocytes TNFalpha + IL-1beta	8.2
Secondary CD8 lymphocyte act	14.3	KU-812 (Basophil) rest	49.3
CD4 lymphocyte none	19.5	KU-812 (Basophil) PMA/ionomycin	74.2
2ry Th1/Th2/Tr1_anti- CD95 CH11	35.8	CCD1106 (Keratinocytes) none	39.5
LAK cells rest	24.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	17.9
LAK cells IL-2	30.8	Liver cirrhosis	3.8
LAK cells IL-2+IL-12	18.6	NCI-H292 none	28.3
LAK cells IL-2+IFN gamma	11.7	NCI-H292 IL-4	52.5

LAK cells IL-2+ IL-18	14.2	NCI-H292 IL-9	62.9
LAK cells PMA/ionomycin	13.3	NCI-H292 IL-13	43.5
NK Cells IL-2 rest	50.7	NCI-H292 IFN gamma	20.6
Two Way MLR 3 day	26.1	HPAEC none	12.9
Two Way MLR 5 day	22.4	HPAEC TNF alpha + IL-1 beta	22.1
Two Way MLR 7 day	23.8	Lung fibroblast none	16.8
PBMC rest	10.0	Lung fibroblast TNF alpha + IL-1 beta	12.9
PBMC PWM	23.2	Lung fibroblast IL-4	13.3
PBMC PHA-L	21.2	Lung fibroblast IL-9	13.5
Ramos (B cell) none	54.3	Lung fibroblast IL-13	14.9
Ramos (B cell) ionomycin	52.1	Lung fibroblast IFN gamma	21.0
B lymphocytes PWM	26.4	Dermal fibroblast CCD1070 rest	36.6
B lymphocytes CD40L and IL-4	27.7	Dermal fibroblast CCD1070 TNF alpha	51.8
EOL-1 dbcAMP	55.1	Dermal fibroblast CCD1070 IL-1 beta	23.7
EOL-1 dbcAMP PMA/ionomycin	49.7	Dermal fibroblast IFN gamma	16.3
Dendritic cells none	49.7	Dermal fibroblast IL-4	29.1
Dendritic cells LPS	20.9	Dermal Fibroblasts rest	13.8
Dendritic cells anti- CD40	41.2	Neutrophils TNFa+LPS	6.7
Monocytes rest	23.5	Neutrophils rest	17.1
Monocytes LPS	36.1	Colon	7.0
Macrophages rest	24.8	Lung	6.9
Macrophages LPS	7.0	Thymus	100.0
HUVEC none	12.7	Kidney	44.1
HUVEC starved	19.6		

Table APD. Panel CNS_1

Tissue Name	Rel. Exp.(%) Ag4070, Run 180912025	Tissue Name	Rel. Exp.(%) Ag4070, Run 180912025
BA4 Control	25.2	BA17 PSP	21.5
BA4 Control2	32.5	BA17 PSP2	6.5
BA4 Alzheimer's2	2.9	Sub Nigra Control	33.7
BA4 Parkinson's	40.1	Sub Nigra Control2	23.2

BA4 Parkinson's2	100.0	Sub Nigra Alzheimer's2	12.1
BA4 Huntington's	29.1	Sub Nigra Parkinson's2	27.7
BA4 Huntington's2	6.1	Sub Nigra Huntington's	49.7
BA4 PSP	17.0	Sub Nigra Huntington's2	28.1
BA4 PSP2	26.6	Sub Nigra PSP2	8.1
BA4 Depression	9.2	Sub Nigra Depression	8.2
BA4 Depression2	10.4	Sub Nigra Depression2	7.2
BA7 Control	37.6	Glob Palladus Control	11.1
BA7 Control2	32.8	Glob Palladus Control2	6.2
BA7 Alzheimer's2	5.1	Glob Palladus Alzheimer's	7.1
BA7 Parkinson's	18.0	Glob Palladus Alzheimer's2	5.5
BA7 Parkinson's2	28.1	Glob Palladus Parkinson's	35.4
BA7 Huntington's	40.9	Glob Palladus Parkinson's2	4.5
BA7 Huntington's2	36.1	Glob Palladus PSP	2.2
BA7 PSP	28.7	Glob Palladus PSP2	10.2
BA7 PSP2	24.7	Glob Palladus Depression	5.7
BA7 Depression	11.7	Temp Pole Control	8.1
BA9 Control	14.2	Temp Pole Control2	18.8
BA9 Control2	39.0	Temp Pole Alzheimer's	3.1
BA9 Alzheimer's	3.9	Temp Pole Alzheimer's2	0.0
BA9 Alzheimer's2	11.0	Temp Pole Parkinson's	19.1
BA9 Parkinson's	16.7	Temp Pole Parkinson's2	14.6
BA9 Parkinson's2	48.0	Temp Pole Huntington's	21.6
BA9 Huntington's	31.9	Temp Pole PSP	2.8
BA9 Huntington's2	21.5	Temp Pole PSP2	1.8
BA9 PSP	10.2	Temp Pole Depression2	5.6
BA9 PSP2	7.0	Cing Gyr Control	41.8
BA9 Depression	6.7	Cing Gyr Control2	13.2
BA9 Depression2	4.6	Cing Gyr Alzheimer's	11.4
BA17 Control	50.3	Cing Gyr Alzheimer's2	6.3
BA17 Control2	58.2	Cing Gyr Parkinson's	8.6
BA17 Alzheimer's2	14.6	Cing Gyr Parkinson's2	24.7
BA17 Parkinson's	38.7	Cing Gyr Huntington's	22.5
BA17 Parkinson's2	53.2	Cing Gyr Huntington's2	12.9
BA17 Huntington's	35.1	Cing Gyr PSP	12.7
BA17 Huntington's2	9.5	Cing Gyr PSP2	1.6
BA17 Depression	9.3	Cing Gyr Depression	5.3
BA17 Depression2	26.8	Cing Gyr Depression2	3.8

Table APE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4070, Run 268390170	Tissue Name	Rel. Exp.(%) Ag4070, Run 268390170
Colon cancer 1	14.3	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	3.5	Bladder cancer NAT 3	0.0
Colon cancer 2	10.1	Bladder cancer NAT 4	8.6
Colon cancer NAT 2	7.5	Adenocarcinoma of the prostate 1	27.7
Colon cancer 3	47.3	Adenocarcinoma of the prostate 2	0.9
Colon cancer NAT 3	14.2	Adenocarcinoma of the prostate 3	16.6
Colon malignant cancer 4	28.3	Adenocarcinoma of the prostate 4	44.4
Colon normal adjacent tissue 4	1.3	Prostate cancer NAT 5	2.8
Lung cancer 1	10.2	Adenocarcinoma of the prostate 6	2.9
Lung NAT 1	0.1	Adenocarcinoma of the prostate 7	3.6
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	2.0
Lung NAT 2	3.5	Adenocarcinoma of the prostate 9	20.0
Squamous cell carcinoma 3	17.9	Prostate cancer NAT 10	0.6
Lung NAT 3	0.3	Kidney cancer 1	9.1
metastatic melanoma 1	15.6	Kidney NAT 1	8.1
Melanoma 2	1.4	Kidney cancer 2	45.7
Melanoma 3	2.2	Kidney NAT 2	17.6
metastatic melanoma 4	39.8	Kidney cancer 3	19.9
metastatic melanoma 5	62.0	Kidney NAT 3	3.1
Bladder cancer 1	2.2	Kidney cancer 4	7.2
Bladder cancer NAT 1	0.0	Kidney NAT 4	1.7
Bladder cancer 2	2.8		

CNS_neurodegeneration_v1.0 Summary: Ag4070 This panel confirms the expression of the CG96581-01 gene at low levels in the brains of an independent group of
5 individuals. However, no differential expression of this gene was detected between

Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Furthermore, low expression of this gene in brain suggests that this gene may play a role in central nervous system disorders such as Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

5 The CG96581-01 gene codes for RP42 homolog. The mouse RP42 gene is expressed in proliferating neuroblasts, whose human orthologs map to susceptibility loci for autism (Mas et al., 2000, Genomics 65: 70-74, PubMed ID:10777668). Therefore, this gene may also play a role in autism and therapeutic modulation of this gene may be useful in the treatment of autism and other neurological disorders.

10 **General_screening_panel_v1.4 Summary:** Ag4070 Results from one experiment with the CG96581-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 4.1D Summary: Ag4070 Highest expression of the CG96581-01 gene is detected in thymus (CT=30.3). This gene is expressed at high to moderate levels in a wide
15 range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in
20 homeostatic processes for these and other cell types and tissues. This pattern suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies,
25 inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

Panel CNS_1 Summary: Ag4070 This panel confirms the expression of the CG96581-01 gene at low levels in the brains. Please see Panel
 CNS_neurodegeneration_v1.0 for a discussion of the potential utility of this gene in
30 treatment of central nervous system disorders.

General oncology screening panel_v_2.4 Summary: Ag4070 Highest expression of the CG96581-01 gene is detected in lung cancer sample (CT=30).

Interestingly, expression of this gene is higher in number of cancer samples including lung, prostate, colon, melanoma and kidney cancers (CTs=30-33) as compared to control normal samples (CTs>35). Therefore, expression of this gene can be used as diagnostic marker for these cancers and therapeutic modulation of this gene can be beneficial in the treatment of these cancers.

AQ. CG96624-01 and CG96624-02: Putative Seven Pass Transmembrane Protein

Expression of gene CG96624-01 and variant CG96624-02 was assessed using the primer-probe sets Ag1372 and Ag4082, described in Tables AQA and AQB. Results of the RTQ-PCR runs are shown in Tables AQC, AQD, AQE and AQF.

Table AQA. Probe Name Ag1372

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ctttgttccccaggtcatc-3'	20	792	359
Probe	TET-5'-cgctggtcagacacattgtaccagt-3'-TAMRA	26	758	360
Reverse	5'-ggctggacaccttcgattac-3'	20	737	361

Table AQB. Probe Name Ag4082

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tgcccttctggttctctac-3'	20	320	362
Probe	TET-5'-cctgcagttcttcaccttgacgtta-3'-TAMRA	26	354	363
Reverse	5'-ttacctgggcaaagtagaggtt-3'	22	382	364

Table AQC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4082, Run 214294986	Tissue Name	Rel. Exp.(%) Ag4082, Run 214294986
AD 1 Hippo	26.8	Control (Path) 3 Temporal Ctx	7.3
AD 2 Hippo	34.2	Control (Path) 4 Temporal Ctx	29.1
AD 3 Hippo	16.6	AD 1 Occipital Ctx	15.4
AD 4 Hippo	9.2	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	95.9	AD 3 Occipital Ctx	15.9
AD 6 Hippo	52.9	AD 4 Occipital Ctx	15.4
Control 2 Hippo	27.7	AD 5 Occipital Ctx	55.1
Control 4 Hippo	15.8	AD 6 Occipital Ctx	22.8
Control (Path) 3 Hippo	8.5	Control 1 Occipital Ctx	4.6

AD 1 Temporal Ctx	23.5	Control 2 Occipital Ctx	89.5
AD 2 Temporal Ctx	33.4	Control 3 Occipital Ctx	18.2
AD 3 Temporal Ctx	11.6	Control 4 Occipital Ctx	9.4
AD 4 Temporal Ctx	18.2	Control (Path) 1 Occipital Ctx	95.9
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	8.5
AD 5 Sup Temporal Ctx	51.4	Control (Path) 3 Occipital Ctx	6.2
AD 6 Inf Temporal Ctx	39.0	Control (Path) 4 Occipital Ctx	15.3
AD 6 Sup Temporal Ctx	36.3	Control 1 Parietal Ctx	8.1
Control 1 Temporal Ctx	11.1	Control 2 Parietal Ctx	42.9
Control 2 Temporal Ctx	66.0	Control 3 Parietal Ctx	18.4
Control 3 Temporal Ctx	19.1	Control (Path) 1 Parietal Ctx	69.3
Control 3 Temporal Ctx	15.8	Control (Path) 2 Parietal Ctx	18.0
Control (Path) 1 Temporal Ctx	58.2	Control (Path) 3 Parietal Ctx	9.7
Control (Path) 2 Temporal Ctx	30.1	Control (Path) 4 Parietal Ctx	36.3

Table AQD. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4082, Run 218897864	Tissue Name	Rel. Exp.(%) Ag4082, Run 218897864
Adipose	1.3	Renal ca. TK-10	8.8
Melanoma* Hs688(A).T	18.8	Bladder	8.2
Melanoma* Hs688(B).T	15.7	Gastric ca. (liver met.) NCI-N87	27.5
Melanoma* M14	38.4	Gastric ca. KATO III	21.8
Melanoma* LOXIMVI	27.0	Colon ca. SW-948	7.2
Melanoma* SK-MEL-5	10.3	Colon ca. SW480	39.8
Squamous cell carcinoma SCC-4	21.0	Colon ca.* (SW480 met) SW620	7.9
Testis Pool	40.3	Colon ca. HT29	9.3
Prostate ca.* (bone met) PC-3	51.4	Colon ca. HCT-116	22.2
Prostate Pool	3.4	Colon ca. CaCo-2	17.2
Placenta	32.3	Colon cancer tissue	7.9
Uterus Pool	1.2	Colon ca. SW1116	6.2
Ovarian ca. OVCAR-3	12.2	Colon ca. Colo-205	5.3
Ovarian ca. SK-OV-3	26.1	Colon ca. SW-48	5.9
Ovarian ca. OVCAR-4	14.2	Colon Pool	9.0
Ovarian ca. OVCAR-5	31.4	Small Intestine Pool	6.2
Ovarian ca. IGROV-1	17.6	Stomach Pool	2.8

Ovarian ca. OVCAR-8	9.3	Bone Marrow Pool	2.4
Ovary	7.2	Fetal Heart	3.6
Breast ca. MCF-7	17.4	Heart Pool	5.1
Breast ca. MDA-MB-231	80.1	Lymph Node Pool	10.1
Breast ca. BT 549	38.2	Fetal Skeletal Muscle	2.0
Breast ca. T47D	100.0	Skeletal Muscle Pool	2.1
Breast ca. MDA-N	37.6	Spleen Pool	3.9
Breast Pool	11.9	Thymus Pool	6.4
Trachea	6.6	CNS cancer (glio/astro) U87-MG	55.9
Lung	0.4	CNS cancer (glio/astro) U-118-MG	38.4
Fetal Lung	12.0	CNS cancer (neuro:met) SK-N-AS	13.8
Lung ca. NCI-N417	7.5	CNS cancer (astro) SF-539	8.8
Lung ca. LX-1	12.0	CNS cancer (astro) SNB-75	32.5
Lung ca. NCI-H146	17.2	CNS cancer (glio) SNB-19	17.1
Lung ca. SHP-77	26.6	CNS cancer (glio) SF-295	36.3
Lung ca. A549	18.2	Brain (Amygdala) Pool	10.2
Lung ca. NCI-H526	20.6	Brain (cerebellum)	38.2
Lung ca. NCI-H23	20.9	Brain (fetal)	23.3
Lung ca. NCI-H460	7.6	Brain (Hippocampus) Pool	10.5
Lung ca. HOP-62	6.4	Cerebral Cortex Pool	15.2
Lung ca. NCI-H522	47.3	Brain (Substantia nigra) Pool	19.8
Liver	2.4	Brain (Thalamus) Pool	16.3
Fetal Liver	12.9	Brain (whole)	17.9
Liver ca. HepG2	3.8	Spinal Cord Pool	9.7
Kidney Pool	9.5	Adrenal Gland	11.2
Fetal Kidney	5.3	Pituitary gland Pool	2.5
Renal ca. 786-0	15.3	Salivary Gland	6.7
Renal ca. A498	8.1	Thyroid (female)	8.8
Renal ca. ACHN	7.3	Pancreatic ca. CAPAN2	21.9
Renal ca. UO-31	17.0	Pancreas Pool	10.7

Table AQE. Panel 1.2

Tissue Name	Rel. Exp.(%) Ag1372, Run 134887810	Tissue Name	Rel. Exp.(%) Ag1372, Run 134887810
Endothelial cells	13.1	Renal ca. 786-0	2.6
Heart (Fetal)	47.6	Renal ca. A498	10.0

Pancreas	0.5	Renal ca. RXF 393	0.7
Pancreatic ca. CAPAN 2	1.3	Renal ca. ACHN	7.5
Adrenal Gland	7.0	Renal ca. UO-31	6.4
Thyroid	2.5	Renal ca. TK-10	5.8
Salivary gland	12.2	Liver	8.5
Pituitary gland	0.2	Liver (fetal)	3.5
Brain (fetal)	1.6	Liver ca. (hepatoblast) HepG2	6.8
Brain (whole)	8.0	Lung	1.4
Brain (amygdala)	0.2	Lung (fetal)	2.1
Brain (cerebellum)	5.3	Lung ca. (small cell) LX-1	7.7
Brain (hippocampus)	32.3	Lung ca. (small cell) NCI-H69	7.6
Brain (thalamus)	20.3	Lung ca. (s.cell var.) SHP-77	1.1
Cerebral Cortex	100.0	Lung ca. (large cell) NCI-H460	13.9
Spinal cord	3.7	Lung ca. (non-sm. cell) A549	9.5
glio/astro U87-MG	15.4	Lung ca. (non-s.cell) NCI-H23	9.7
glio/astro U-118-MG	10.1	Lung ca. (non-s.cell) HOP-62	8.7
astrocytoma SW1783	3.8	Lung ca. (non-s.cl) NCI-H522	71.7
neuro*; met SK-N-AS	14.8	Lung ca. (squam.) SW 900	7.3
astrocytoma SF-539	2.8	Lung ca. (squam.) NCI-H596	19.6
astrocytoma SNB-75	2.5	Mammary gland	3.4
glioma SNB-19	12.6	Breast ca.* (pl.ef) MCF-7	2.7
glioma U251	0.0	Breast ca.* (pl.ef) MDA-MB-231	6.4
glioma SF-295	19.5	Breast ca.* (pl. ef) T47D	11.6
Heart	15.0	Breast ca. BT-549	3.1
Skeletal Muscle	4.5	Breast ca. MDA-N	50.7
Bone marrow	2.5	Ovary	22.8
Thymus	1.7	Ovarian ca. OVCAR-3	0.0
Spleen	3.6	Ovarian ca. OVCAR-4	19.6
Lymph node	1.7	Ovarian ca. OVCAR-5	25.2
Colorectal Tissue	2.1	Ovarian ca. OVCAR-8	23.7
Stomach	9.7	Ovarian ca. IGROV-1	6.9
Small intestine	14.5	Ovarian ca. (ascites) SK-OV-3	10.4
Colon ca. SW480	2.6	Uterus	3.7
Colon ca.* SW620 (SW480 met)	1.5	Placenta	32.8
Colon ca. HT29	1.2	Prostate	23.0
Colon ca. HCT-116	1.0	Prostate ca.* (bone met) PC-3	52.5

Colon ca. CaCo-2	3.7	Testis	10.4
Colon ca. Tissue (ODO3866)	1.0	Melanoma Hs688(A).T	3.8
Colon ca. HCC-2998	7.1	Melanoma* (met) Hs688(B).T	2.5
Gastric ca.* (liver met) NCI-N87	7.5	Melanoma UACC-62	29.3
Bladder	9.5	Melanoma M14	9.3
Trachea	1.3	Melanoma LOX IMVI	2.6
Kidney	50.3	Melanoma* (met) SK-MEL-5	4.4
Kidney (fetal)	3.1		

Table AQF. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4082, Run 171809988	Tissue Name	Rel. Exp.(%) Ag4082, Run 171809988
Secondary Th1 act	44.1	HUVEC IL-1beta	29.3
Secondary Th2 act	55.5	HUVEC IFN gamma	24.7
Secondary Tr1 act	45.1	HUVEC TNF alpha + IFN gamma	29.5
Secondary Th1 rest	9.2	HUVEC TNF alpha + IL4	47.0
Secondary Th2 rest	10.8	HUVEC IL-11	13.9
Secondary Tr1 rest	10.8	Lung Microvascular EC none	58.2
Primary Th1 act	29.9	Lung Microvascular EC TNFalpha + IL-1beta	66.4
Primary Th2 act	47.6	Microvascular Dermal EC none	20.2
Primary Tr1 act	45.7	Microvascular Dermal EC TNFalpha + IL-1beta	37.1
Primary Th1 rest	9.2	Bronchial epithelium TNFalpha + IL1beta	31.2
Primary Th2 rest	5.8	Small airway epithelium none	9.2
Primary Tr1 rest	14.7	Small airway epithelium TNFalpha + IL-1beta	17.9
CD45RA CD4 lymphocyte act	40.6	Coronary artery SMC rest	16.0
CD45RO CD4 lymphocyte act	45.1	Coronary artery SMC TNFalpha + IL-1beta	18.3
CD8 lymphocyte act	51.8	Astrocytes rest	25.3
Secondary CD8 lymphocyte rest	60.7	Astrocytes TNFalpha + IL-1beta	27.0
Secondary CD8 lymphocyte act	29.5	KU-812 (Basophil) rest	13.2
CD4 lymphocyte none	6.9	KU-812 (Basophil)	37.4

		PMA/ionomycin	
2ry Th1/Th2/Tr1_anti-CD95 CH11	19.3	CCD1106 (Keratinocytes) none	31.2
LAK cells rest	31.4	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	23.0
LAK cells IL-2	28.9	Liver cirrhosis	4.5
LAK cells IL-2+IL-12	28.1	NCI-H292 none	40.3
LAK cells IL-2+IFN gamma	14.7	NCI-H292 IL-4	0.2
LAK cells IL-2+ IL-18	25.9	NCI-H292 IL-9	100.0
LAK cells PMA/ionomycin	27.4	NCI-H292 IL-13	83.5
NK Cells IL-2 rest	55.9	NCI-H292 IFN gamma	99.3
Two Way MLR 3 day	45.1	HPAEC none	16.5
Two Way MLR 5 day	36.9	HPAEC TNF alpha + IL-1 beta	68.3
Two Way MLR 7 day	17.9	Lung fibroblast none	27.0
PBMC rest	9.6	Lung fibroblast TNF alpha + IL-1 beta	29.7
PBMC PWM	49.3	Lung fibroblast IL-4	29.9
PBMC PHA-L	25.5	Lung fibroblast IL-9	28.3
Ramos (B cell) none	35.4	Lung fibroblast IL-13	26.4
Ramos (B cell) ionomycin	31.6	Lung fibroblast IFN gamma	32.8
B lymphocytes PWM	34.4	Dermal fibroblast CCD1070 rest	23.2
B lymphocytes CD40L and IL-4	41.2	Dermal fibroblast CCD1070 TNF alpha	40.3
EOL-1 dbcAMP	12.8	Dermal fibroblast CCD1070 IL-1 beta	19.6
EOL-1 dbcAMP PMA/ionomycin	15.8	Dermal fibroblast IFN gamma	28.1
Dendritic cells none	24.5	Dermal fibroblast IL-4	30.8
Dendritic cells LPS	37.6	Dermal Fibroblasts rest	24.8
Dendritic cells anti-CD40	32.1	Neutrophils TNFa+LPS	3.5
Monocytes rest	11.7	Neutrophils rest	8.0
Monocytes LPS	33.7	Colon	9.0
Macrophages rest	72.2	Lung	14.9
Macrophages LPS	25.2	Thymus	11.3
HUVEC none	15.7	Kidney	24.0
HUVEC starved	23.3		

CNS_neurodegeneration_v1.0 Summary: Ag4082 This panel does not show differential expression of the CG96624-01 gene in Alzheimer's disease. However, this

expression profile confirms the presence of this gene in the brain. Please see Panel 1.4 for discussion of utility of this gene in the central nervous system.

General_screening_panel_v1.4 Summary: Ag4082 Highest expression of the CG96624-01 gene is seen in a breast cancer cell line (CT=25.8). Significant levels of expression are also seen in a cluster of samples derived from breast, brain, colon, lung, ovarian, prostate and melanoma cancer cell lines.

In addition, this gene is expressed at much higher levels in fetal lung (CT=28.9) when compared to expression in the adult counterpart (CT=36.7). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue.

The ubiquitous expression of this gene in this panel and the higher levels of expression in cancer cell lines and some fetal tissues suggest a role for this gene product in cell survival and proliferation.

Among tissues with metabolic function, this gene is expressed at moderate levels in pituitary, adrenal gland, pancreas, thyroid, and adult and fetal skeletal muscle, heart, and liver and at low but significant levels in adipose. This widespread expression among these tissues suggests that this gene product may play a role in normal neuroendocrine and metabolic function and that dysregulated expression of this gene may contribute to neuroendocrine disorders or metabolic diseases, such as obesity and diabetes.

This gene is also expressed at high to moderate levels in the CNS, including the hippocampus, thalamus, substantia nigra, amygdala, cerebellum and cerebral cortex. Therefore, therapeutic modulation of the expression or function of this gene may be useful in the treatment of neurologic disorders, such as Alzheimer's disease, Parkinson's disease, schizophrenia, multiple sclerosis, stroke and epilepsy.

Panel 1.2 Summary: Ag4082 Highest expression of the CG96624-01 gene is seen in the cerebral cortex (CT=23.8). Expression of this gene is ubiquitous in this panel, with prominent levels of expression in lung and breast cancer cell lines and normal kidney (CT=24.8). Thus, expression of this gene could be used as a marker of lung and breast cancer and to differentiate kidney from fetal kidney (CT=28.8). Please see Panel 1.4 for further discussion of utility of this gene.

Panel 4.1D Summary: Ag4082 Highest expression of the CG96624-01 gene is seen in IL-9 treated NCI-H292 cells (CT=29.6). In addition, this gene is expressed at moderate levels in a wide range of cell types of significance in the immune response in

health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

AR. CG96747-01: Voltage-Dependent Calcium Channel Gamma-3 Subunit - Like Protein

Expression of gene CG96747-01 was assessed using the primer-probe set Ag4076, described in Table ARA. Results of the RTQ-PCR runs are shown in Tables ARB and ARC.

Table ARA. Probe Name Ag4076

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agcgtgtctctgctgcttt-3'	19	421	365
Probe	TET-5'-tcaccggtgctacttctgctg-3'-TAMRA	23	440	366
Reverse	5'-aggtgcgagtagctgatgtaga-3'	22	494	367

Table ARB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4076, Run 214294983	Tissue Name	Rel. Exp.(%) Ag4076, Run 214294983
AD 1 Hippo	15.1	Control (Path) 3 Temporal Ctx	0.0
AD 2 Hippo	71.2	Control (Path) 4 Temporal Ctx	13.3
AD 3 Hippo	6.4	AD 1 Occipital Ctx	9.1
AD 4 Hippo	7.1	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	53.2	AD 3 Occipital Ctx	2.4

AD 6 Hippo	15.2	AD 4 Occipital Ctx	23.2
Control 2 Hippo	34.2	AD 5 Occipital Ctx	14.0
Control 4 Hippo	29.3	AD 6 Occipital Ctx	49.0
Control (Path) 3 Hippo	3.5	Control 1 Occipital Ctx	11.7
AD 1 Temporal Ctx	17.0	Control 2 Occipital Ctx	74.7
AD 2 Temporal Ctx	43.8	Control 3 Occipital Ctx	9.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	11.3
AD 4 Temporal Ctx	13.6	Control (Path) 1 Occipital Ctx	97.9
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	8.8
AD 5 Sup Temporal Ctx	48.0	Control (Path) 3 Occipital Ctx	0.9
AD 6 Inf Temporal Ctx	27.9	Control (Path) 4 Occipital Ctx	4.2
AD 6 Sup Temporal Ctx	17.9	Control 1 Parietal Ctx	32.8
Control 1 Temporal Ctx	16.7	Control 2 Parietal Ctx	48.3
Control 2 Temporal Ctx	50.3	Control 3 Parietal Ctx	14.0
Control 3 Temporal Ctx	8.8	Control (Path) 1 Parietal Ctx	42.9
Control 4 Temporal Ctx	25.9	Control (Path) 2 Parietal Ctx	17.4
Control (Path) 1 Temporal Ctx	27.0	Control (Path) 3 Parietal Ctx	2.0
Control (Path) 2 Temporal Ctx	7.0	Control (Path) 4 Parietal Ctx	5.8

Table ARC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4076, Run 218903895	Tissue Name	Rel. Exp.(%) Ag4076, Run 218903895
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	0.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.0
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0

Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.0	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.0
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.0
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.0	Thymus Pool	0.0
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	1.1	CNS cancer (neuro;met) SK-N-AS	0.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	54.7
Lung ca. NCI-H526	0.0	Brain (cerebellum)	37.1
Lung ca. NCI-H23	0.0	Brain (fetal)	0.0
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	41.2
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	38.2
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	100.0
Liver	0.0	Brain (Thalamus) Pool	45.7
Fetal Liver	0.0	Brain (whole)	18.2
Liver ca. HepG2	0.0	Spinal Cord Pool	67.8
Kidney Pool	0.0	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	0.0

CNS_neurodegeneration_v1.0 Summary: Ag4076 This panel confirms the expression of the CG96747-01 gene at low levels in the brains of an independent group of

individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

- 5 **General_screening_panel_v1.4 Summary:** Ag4076 Highest expression of the CG96747-01 gene is detected in substantia nigra of brain (CT=30.1). Interestingly expression of this gene is exclusive to brain regions examined. Therefore, expression of this gene can be used to distinguish brain from other other samples used in this panel. Furthermore, therapeutic modulation of this gene through use of small molecule target
10 may be useful in the treatment of neurological disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

- The CG96747-01 gene codes for a protein belonging to PMP-22/EMP/MP20/Claudin family. Proteins belonging to this family are small integral membrane glycoproteins which are evolutionarily related including eye lens specific
15 membrane protein 20 (MP20 or MP19); epithelial membrane protein-1 (EMP-1/TMP), epithelial membrane protein-2 (EMP-2), and peripheral myelin protein 22 (PMP-22). PMP-22 plays a role both in myelination and in cell proliferation. Mutations affecting PMP-22 are associated with hereditary motor and sensory neuropathies such as Charcot-Marie-Tooth disease type 1A (CMT-1A) in human or the trembler phenotype in mice
20 (Jetten AM, Suter U, 2000, Prog Nucleic Acid Res Mol Biol 64:97-129, PMID: 10697408; PFAM: IPR004031). Thus, the CG96747-01 gene product may also play a role in hereditary motor and sensory neuropathies such as CMT-1A and therapeutic modulation of this gene product may be useful in the treatment of this disease.

- Panel 4.1D Summary:** Ag4076 Expression of the CG96747-01 gene is
25 low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4076 Expression of the CG96747-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

30 **AS. CG97462-01: Prohibitin**

 Expression of gene CG97462-01 was assessed using the primer-probe set Ag4111, described in Table ASA.

Table ASA. Probe Name Ag4111

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggctctccaggagcaattac-3'	20	432	368
Probe	TET-5'-ttacagagcgagcagccaccttg-3'-TAMRA	24	452	369
Reverse	5'-atgtgtcaaggacacgtcatc-3'	21	487	370

CNS_neurodegeneration_v1.0 Summary: Ag4111 Expression of the CG97462-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

5 **General_screening_panel_v1.4 Summary:** Ag4111 Expression of the CG97462-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4111 Expression of the CG97462-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

10 **General oncology screening panel_v_2.4 Summary:** Ag4111 Expression of the CG97462-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

AT. CG97472-01: Glucose Transporter

15 Expression of gene CG97472-01 was assessed using the primer-probe set Ag4113, described in Table ATA. Results of the RTQ-PCR runs are shown in Tables ATB, ATC, ATD and ATE.

Table ATA. Probe Name Ag4113

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agctcctttttctgttggaactgt-3'	22	235	371
Probe	TET-5'-tcaacagctttgacaggcgtaattca-3'-TAMRA	26	260	372
Reverse	5'-cagccaacagggtgacagtaag-3'	22	289	373

Table ATB. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4113, Run 219543072	Tissue Name	Rel. Exp.(%) Ag4113, Run 219543072
Adipose	3.4	Renal ca. TK-10	16.5
Melanoma* Hs688(A).T	2.4	Bladder	11.4

Melanoma* Hs688(B).T	1.0	Gastric ca. (liver met.) NCI-N87	23.3
Melanoma* M14	2.1	Gastric ca. KATO III	2.0
Melanoma* LOXIMVI	0.8	Colon ca. SW-948	0.2
Melanoma* SK-MEL-5	2.3	Colon ca. SW480	5.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	3.9
Testis Pool	5.5	Colon ca. HT29	1.6
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	1.3
Prostate Pool	7.4	Colon ca. CaCo-2	2.1
Placenta	1.0	Colon cancer tissue	2.8
Uterus Pool	6.2	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	4.3	Colon ca. Colo-205	0.6
Ovarian ca. SK-OV-3	3.5	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	37.1
Ovarian ca. OVCAR-5	28.7	Small Intestine Pool	71.2
Ovarian ca. IGROV-1	2.0	Stomach Pool	25.3
Ovarian ca. OVCAR-8	1.3	Bone Marrow Pool	17.8
Ovary	14.3	Fetal Heart	3.6
Breast ca. MCF-7	2.0	Heart Pool	11.9
Breast ca. MDA-MB-231	6.0	Lymph Node Pool	36.1
Breast ca. BT 549	6.3	Fetal Skeletal Muscle	3.9
Breast ca. T47D	25.7	Skeletal Muscle Pool	2.8
Breast ca. MDA-N	3.1	Spleen Pool	5.5
Breast Pool	35.1	Thymus Pool	15.6
Trachea	8.7	CNS cancer (glio/astro) U87-MG	8.8
Lung	24.8	CNS cancer (glio/astro) U-118-MG	1.7
Fetal Lung	16.5	CNS cancer (neuro;met) SK-N-AS	5.1
Lung ca. NCI-N417	0.9	CNS cancer (astro) SF-539	2.0
Lung ca. LX-1	8.9	CNS cancer (astro) SNB-75	4.0
Lung ca. NCI-H146	0.2	CNS cancer (glio) SNB-19	4.4
Lung ca. SHP-77	1.7	CNS cancer (glio) SF-295	7.6
Lung ca. A549	4.4	Brain (Amygdala) Pool	4.1
Lung ca. NCI-H526	1.9	Brain (cerebellum)	3.6
Lung ca. NCI-H23	9.9	Brain (fetal)	8.7
Lung ca. NCI-H460	5.9	Brain (Hippocampus) Pool	7.7
Lung ca. HOP-62	5.3	Cerebral Cortex Pool	8.7

Lung ca. NCI-H522	3.8	Brain (Substantia nigra) Pool	7.9
Liver	0.4	Brain (Thalamus) Pool	12.7
Fetal Liver	1.7	Brain (whole)	6.0
Liver ca. HepG2	0.0	Spinal Cord Pool	7.0
Kidney Pool	100.0	Adrenal Gland	11.0
Fetal Kidney	12.9	Pituitary gland Pool	6.3
Renal ca. 786-0	3.6	Salivary Gland	2.4
Renal ca. A498	0.0	Thyroid (female)	2.6
Renal ca. ACHN	2.7	Pancreatic ca. CAPAN2	9.8
Renal ca. UO-31	0.5	Pancreas Pool	30.4

Table ATC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4113, Run 172792545	Tissue Name	Rel. Exp.(%) Ag4113, Run 172792545
Secondary Th1 act	0.0	HUVEC IL-1beta	2.8
Secondary Th2 act	3.8	HUVEC IFN gamma	41.2
Secondary Tr1 act	1.8	HUVEC TNF alpha + IFN gamma	9.7
Secondary Th1 rest	12.0	HUVEC TNF alpha + IL4	1.4
Secondary Th2 rest	21.8	HUVEC IL-11	6.1
Secondary Tr1 rest	24.7	Lung Microvascular EC none	21.8
Primary Th1 act	5.3	Lung Microvascular EC TNFalpha + IL-1beta	31.6
Primary Th2 act	0.0	Microvascular Dermal EC none	4.0
Primary Tr1 act	0.3	Microvascular Dermal EC TNFalpha + IL-1beta	14.2
Primary Th1 rest	8.7	Bronchial epithelium TNFalpha + IL1beta	5.7
Primary Th2 rest	6.4	Small airway epithelium none	0.0
Primary Tr1 rest	9.6	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	10.3	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	24.5	Coronary artery SMC TNFalpha + IL-1beta	3.3
CD8 lymphocyte act	4.4	Astrocytes rest	1.3
Secondary CD8 lymphocyte rest	0.6	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	23.8

CD4 lymphocyte none	5.3	KU-812 (Basophil) PMA/ionomycin	23.0
2ry Th1/Th2/Tr1_anti- CD95 CH11	27.7	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	12.9	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.2
LAK cells IL-2	4.7	Liver cirrhosis	3.6
LAK cells IL-2+IL-12	13.2	NCI-H292 none	12.9
LAK cells IL-2+IFN gamma	5.5	NCI-H292 IL-4	4.5
LAK cells IL-2+ IL-18	10.4	NCI-H292 IL-9	17.3
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	12.3
NK Cells IL-2 rest	17.6	NCI-H292 IFN gamma	21.9
Two Way MLR 3 day	20.3	HPAEC none	14.8
Two Way MLR 5 day	10.9	HPAEC TNF alpha + IL-1 beta	17.2
Two Way MLR 7 day	0.9	Lung fibroblast none	1.5
PBMC rest	2.5	Lung fibroblast TNF alpha + IL-1 beta	5.4
PBMC PWM	7.3	Lung fibroblast IL-4	4.0
PBMC PHA-L	10.7	Lung fibroblast IL-9	3.5
Ramos (B cell) none	5.7	Lung fibroblast IL-13	10.2
Ramos (B cell) ionomycin	1.5	Lung fibroblast IFN gamma	4.8
B lymphocytes PWM	1.5	Dermal fibroblast CCD1070 rest	12.2
B lymphocytes CD40L and IL-4	0.9	Dermal fibroblast CCD1070 TNF alpha	12.8
EOL-1 dbcAMP	6.0	Dermal fibroblast CCD1070 IL-1 beta	0.4
EOL-1 dbcAMP PMA/ionomycin	3.1	Dermal fibroblast IFN gamma	2.3
Dendritic cells none	9.6	Dermal fibroblast IL-4	2.5
Dendritic cells LPS	21.5	Dermal Fibroblasts rest	11.5
Dendritic cells anti-CD40	5.3	Neutrophils TNFa+LPS	11.1
Monocytes rest	14.0	Neutrophils rest	92.7
Monocytes LPS	88.9	Colon	24.7
Macrophages rest	8.7	Lung	29.7
Macrophages LPS	6.5	Thymus	56.6
HUVEC none	3.7	Kidney	100.0
HUVEC starved	3.1		

Table ATD, Panel 5D

Tissue Name	Rel. Exp.(%)	Tissue Name	Rel.
-------------	--------------	-------------	------

	Ag4113, Run 172784076		Exp.(%) Ag4113, Run 172784076
97457_Patient-02go_adipose	71.2	94709_Donor 2 AM - A_adipose	0.0
97476_Patient-07sk_skeletal muscle	13.2	94710_Donor 2 AM - B_adipose	0.0
97477_Patient-07ut_uterus	35.4	94711_Donor 2 AM - C_adipose	0.0
97478_Patient-07pl_placenta	0.0	94712_Donor 2 AD - A_adipose	36.3
97481_Patient-08sk_skeletal muscle	40.6	94713_Donor 2 AD - B_adipose	12.4
97482_Patient-08ut_uterus	9.4	94714_Donor 2 AD - C_adipose	29.3
97483_Patient-08pl_placenta	22.5	94742_Donor 3 U - A_Mesenchymal Stem Cells	0.0
97486_Patient-09sk_skeletal muscle	0.0	94743_Donor 3 U - B_Mesenchymal Stem Cells	6.8
97487_Patient-09ut_uterus	30.8	94730_Donor 3 AM - A_adipose	0.0
97488_Patient-09pl_placenta	0.0	94731_Donor 3 AM - B_adipose	0.0
97492_Patient-10ut_uterus	32.5	94732_Donor 3 AM - C_adipose	0.0
97493_Patient-10pl_placenta	29.3	94733_Donor 3 AD - A_adipose	10.6
97495_Patient-11go_adipose	100.0	94734_Donor 3 AD - B_adipose	11.6
97496_Patient-11sk_skeletal muscle	23.2	94735_Donor 3 AD - C_adipose	0.0
97497_Patient-11ut_uterus	26.4	77138_Liver_HepG2untreated	0.0
97498_Patient-11pl_placenta	12.9	73556_Heart_Cardiac stromal cells (primary)	15.7
97500_Patient-12go_adipose	97.9	81735_Small Intestine	44.1
97501_Patient-12sk_skeletal muscle	82.9	72409_Kidney Proximal Convoluted Tubule	0.0
97502_Patient-12ut_uterus	15.7	82685_Small intestine_Duodenum	21.6
97503_Patient-12pl_placenta	25.2	90650_Adrenal_Adrenocortical adenoma	0.0
94721_Donor 2 U - A_Mesenchymal Stem Cells	0.0	72410_Kidney_HRCE	37.1
94722_Donor 2 U - B_Mesenchymal Stem Cells	0.0	72411_Kidney_HRE	6.9

94723_Donor 2 U - C_Mesenchymal Stem Cells	14.7	73139_Uterus_Uterine smooth muscle cells	0.0
--	------	---	-----

Table ATE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4113, Run 268389993	Tissue Name	Rel. Exp.(%) Ag4113. Run 268389993
Colon cancer 1	7.7	Bladder cancer NAT 2	1.0
Colon cancer NAT 1	1.8	Bladder cancer NAT 3	0.0
Colon cancer 2	2.6	Bladder cancer NAT 4	3.1
Colon cancer NAT 2	2.0	Adenocarcinoma of the prostate 1	70.2
Colon cancer 3	7.6	Adenocarcinoma of the prostate 2	3.7
Colon cancer NAT 3	15.4	Adenocarcinoma of the prostate 3	11.0
Colon malignant cancer 4	7.7	Adenocarcinoma of the prostate 4	16.8
Colon normal adjacent tissue 4	2.6	Prostate cancer NAT 5	1.6
Lung cancer 1	6.1	Adenocarcinoma of the prostate 6	3.1
Lung NAT 1	1.5	Adenocarcinoma of the prostate 7	7.3
Lung cancer 2	9.1	Adenocarcinoma of the prostate 8	2.9
Lung NAT 2	4.9	Adenocarcinoma of the prostate 9	37.4
Squamous cell carcinoma 3	10.2	Prostate cancer NAT 10	3.3
Lung NAT 3	0.6	Kidney cancer 1	35.6
metastatic melanoma 1	46.3	KidneyNAT 1	8.9
Melanoma 2	0.5	Kidney cancer 2	46.3
Melanoma 3	0.9	Kidney NAT 2	9.2
metastatic melanoma 4	90.8	Kidney cancer 3	29.1
metastatic melanoma 5	100.0	Kidney NAT 3	5.6
Bladder cancer 1	2.2	Kidney cancer 4	5.8
Bladder cancer NAT 1	0.0	Kidney NAT 4	5.3
Bladder cancer 2	6.0		

CNS_neurodegeneration_v1.0 Summary: Ag4113 Results from one experiment with the CG97472-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

5

General_screening_panel_v1.4 Summary: Ag4113 Highest expression of the CG97472-01 gene is detected in kidney (CT=29.5). Therefore, therapeutic modulation of this gene may be useful in the treatment of kidney related disease including lupus and glomerulonephritis.

Among tissues with metabolic or endocrine function, this gene is expressed at low to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at moderate to low levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Low level expression of this gene is also seen in cluster of ovarian cancer, breast, lung, renal, colon and CNS cancer. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

Panel 4.1D Summary: Ag4113 Highest expression of the CG97472-01 gene is detected in kidney (CT=31). This gene is expressed at low to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

Panel 5D Summary: Ag4113 Highest expression of the CG97472-01 gene is detected in adipose (CT=34.4). Low level of expression of this gene is seen exclusively in adipose and skeletal muscle samples. Therefore, expression of this gene can be used to

distinguish these samples from other samples used in this panel. Please see Panel 1.4 for a discussion of the potential utility of this gene

General oncology screening panel_v_2.4 Summary: Ag4113 Highest expression of the CG97472-01 gene is detected in metastatic melanoma sample (CT=31).

- 5 Interestingly, significant expression of this gene is seen in number of cancer samples including kidney, metastatic melanoma, prostate adenocarcinoma and lung cancer. Therefore, expression of this gene may be used as diagnostic marker for these cancers and therapeutic modulation of this gene can be useful in the treatment of these cancers.

10 AU. CG97528-01: Guanylate Binding Protein

Expression of gene CG97528-01 was assessed using the primer-probe set Ag4107, described in Table AUA. Results of the RTQ-PCR runs are shown in Tables AUB and AUC.

Table AUA. Probe Name Ag4107

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aacaggcccgagtactaaagg-3'	21	1814	374
Probe	TET-5'-tgccaagggtgaaagtacccaacttca-3'-TAMRA	26	1840	375
Reverse	5'-tcagggtcttctgtagcttttg-3'	22	1876	376

15 **Table AUB. General screening_panel_v1.4**

Tissue Name	Rel. Exp.(%) Ag4107, Run 219447017	Tissue Name	Rel. Exp.(%) Ag4107, Run 219447017
Adipose	5.5	Renal ca. TK-10	3.3
Melanoma* Hs688(A).T	12.5	Bladder	19.5
Melanoma* Hs688(B).T	11.0	Gastric ca. (liver met.) NCI-N87	100.0
Melanoma* M14	10.8	Gastric ca. KATO III	24.0
Melanoma* LOXIMVI	3.4	Colon ca. SW-948	4.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	6.1
Squamous cell carcinoma SCC-4	4.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	1.3	Colon ca. HT29	4.0
Prostate ca.* (bone met) PC-3	11.6	Colon ca. HCT-116	0.6
Prostate Pool	4.9	Colon ca. CaCo-2	0.4

Placenta	0.7	Colon cancer tissue	11.8
Uterus Pool	3.0	Colon ca. SW1116	0.5
Ovarian ca. OVCAR-3	3.8	Colon ca. Colo-205	0.6
Ovarian ca. SK-OV-3	6.9	Colon ca. SW-48	0.7
Ovarian ca. OVCAR-4	2.8	Colon Pool	7.2
Ovarian ca. OVCAR-5	22.8	Small Intestine Pool	5.3
Ovarian ca. IGROV-1	1.9	Stomach Pool	2.8
Ovarian ca. OVCAR-8	1.4	Bone Marrow Pool	3.3
Ovary	4.3	Fetal Heart	1.1
Breast ca. MCF-7	1.2	Heart Pool	2.5
Breast ca. MDA-MB-231	40.9	Lymph Node Pool	6.8
Breast ca. BT 549	5.8	Fetal Skeletal Muscle	0.7
Breast ca. T47D	29.3	Skeletal Muscle Pool	2.7
Breast ca. MDA-N	1.1	Spleen Pool	12.2
Breast Pool	5.6	Thymus Pool	5.2
Trachea	5.2	CNS cancer (glio/astro) U87-MG	0.7
Lung	0.8	CNS cancer (glio/astro) U-118-MG	40.6
Fetal Lung	14.4	CNS cancer (neuro;met) SK-N-AS	2.6
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	8.2
Lung ca. LX-1	0.9	CNS cancer (astro) SNB-75	17.9
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	1.5
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	40.1
Lung ca. A549	0.5	Brain (Amygdala) Pool	0.4
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.1
Lung ca. NCI-H23	0.2	Brain (fetal)	0.3
Lung ca. NCI-H460	0.2	Brain (Hippocampus) Pool	0.7
Lung ca. HOP-62	4.0	Cerebral Cortex Pool	0.8
Lung ca. NCI-H522	0.2	Brain (Substantia nigra) Pool	0.3
Liver	0.2	Brain (Thalamus) Pool	1.1
Fetal Liver	0.7	Brain (whole)	0.3
Liver ca. HepG2	0.1	Spinal Cord Pool	2.1
Kidney Pool	9.6	Adrenal Gland	2.1
Fetal Kidney	4.7	Pituitary gland Pool	0.8
Renal ca. 786-0	13.0	Salivary Gland	0.3
Renal ca. A498	3.8	Thyroid (female)	3.2
Renal ca. ACHN	2.8	Pancreatic ca. CAPAN2	1.9
Renal ca. UO-31	6.0	Pancreas Pool	8.7

Table AUC. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4107, Run 268623659	Tissue Name	Rel. Exp.(%) Ag4107, Run 268623659
Colon cancer 1	22.4	Bladder cancer NAT 2	0.8
Colon NAT 1	13.9	Bladder cancer NAT 3	1.3
Colon cancer 2	36.1	Bladder cancer NAT 4	2.4
Colon cancer NAT 2	19.5	Adenocarcinoma of the prostate 1	69.7
Colon cancer 3	76.3	Adenocarcinoma of the prostate 2	6.5
Colon cancer NAT 3	66.4	Adenocarcinoma of the prostate 3	19.8
Colon malignant cancer 4	64.6	Adenocarcinoma of the prostate 4	51.1
Colon normal adjacent tissue 4	28.3	Prostate cancer NAT 5	7.1
Lung cancer 1	23.3	Adenocarcinoma of the prostate 6	4.8
Lung NAT 1	5.8	Adenocarcinoma of the prostate 7	6.4
Lung cancer 2	11.3	Adenocarcinoma of the prostate 8	1.6
Lung NAT 2	5.4	Adenocarcinoma of the prostate 9	28.5
Squamous cell carcinoma 3	21.9	Prostate cancer NAT 10	1.7
Lung NAT 3	2.7	Kidney cancer 1	50.0
metastatic melanoma 1	27.5	Kidney NAT 1	18.7
Melanoma 2	0.8	Kidney cancer 2	100.0
Melanoma 3	2.4	Kidney NAT 2	43.2
metastatic melanoma 4	50.7	Kidney cancer 3	94.0
metastatic melanoma 5	47.6	Kidney NAT 3	11.9
Bladder cancer 1	3.3	Kidney cancer 4	16.5
Bladder cancer NAT 1	0.0	Kidney NAT 4	3.1
Bladder cancer 2	9.2		

CNS_neurodegeneration_v1.0 Summary: Ag4107 Results from one experiment with the CG97528-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

- 5 **General_screening_panel_v1.4 Summary:** Ag4107 Highest expression of the CG97528-01 gene is detected in gastric cancer cell line (CT=25). High expression of this gene is also seen in cluster of cancer cell lines including CNS, colon, renal, breast, ovarian, prostate, squamous cell carcinoma and melanoma. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

Significant expression is also detected in fetal and adult lung. Interestingly, this gene is expressed at much higher levels in fetal (CT = 27.8) when compared to adult lung (CT=32). This observation suggests that expression of this gene can be used to distinguish fetal from adult lung. In addition, the relative overexpression of this gene in fetal lung suggests that the protein product may enhance lung growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of lung related diseases.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at moderate to low levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4107 Results from one experiment with the CG97528-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General oncology screening panel_v_2.4 Summary: Ag4107 Highest expression of the CG97528-01 gene is detected in kidney cancer (CT=27). Interestingly, expression of this gene is higher in number of cancer samples including kidney, adenocarcinoma of the prostate, melanoma, lung, and colon cancers. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

AV. CG97629-01: Cell Division Protein Kinase 7

Expression of gene CG97629-01 was assessed using the primer-probe set Ag4122, described in Table AVA.

Table AVA. Probe Name Ag4122

Primers	Sequences	Length	Start Position	SEQ ID No
---------	-----------	--------	----------------	-----------

Forward	5'-gggacattttgctacagcctat-3'	22	142	377
Probe	TET-5'-agaacaccaaccaaatacgtcaccatt-3'-TAMRA	26	177	378
Reverse	5'-agcttctgacctgtgtccaa-3'	20	216	379

CNS_neurodegeneration_v1.0 Summary: Ag4122 Expression of the CG97629-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General_screening_panel_v1.4 Summary: Ag4122 Expression of the CG97629-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4122 Expression of the CG97629-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel CNS_1 Summary: Ag4122 Expression of the CG97629-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4122 Expression of the CG97629-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

15 **AW. CG97648-01: G Protein-Coupled Receptor Kinase GRK7**

Expression of gene CG97648-01 was assessed using the primer-probe sets Ag3046 and Ag4125, described in Tables AWA and AWB. Results of the RTQ-PCR runs are shown in Tables AWC, AWD, AWE and AWF.

Table AWA. Probe Name Ag3046

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gaagcaaagaactctgcaagac-3'	22	1281	380
Probe	TET-5'-ttccagcatgataacttcacagagga-3'-TAMRA	26	1312	381
Reverse	5'-gagcctgcaaataatcttttgc-3'	22	1338	382

20

Table AWB. Probe Name Ag4125

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gaagcaaagaactctgcaagac-3'	22	1281	383
Probe	TET-5'-ttccagcatgataacttcacagagga-3'-TAMRA	26	1312	384
Reverse	5'-gagcctgcaaataatcttttgc-3'	22	1338	385

Table AWC. Panel 1.3D

Tissue Name	Rel. Exp.(%) Ag3046, Run 165533236	Rel. Exp.(%) Ag3046, Run 165724481	Tissue Name	Rel. Exp.(%) Ag3046, Run 165533236	Rel. Exp.(%) Ag3046, Run 165724481
Liver adenocarcinoma	27.5	0.0	Kidney (fetal)	0.0	0.0
Pancreas	0.0	72.2	Renal ca. 786-0	0.0	0.0
Pancreatic ca. CAPAN 2	0.0	0.0	Renal ca. A498	43.2	0.0
Adrenal gland	0.0	0.0	Renal ca. RXF 393	0.0	0.0
Thyroid	0.0	0.0	Renal ca. ACHN	0.0	0.0
Salivary gland	0.0	25.5	Renal ca. UO-31	0.0	0.0
Pituitary gland	0.0	0.0	Renal ca. TK-10	0.0	0.0
Brain (fetal)	38.4	0.0	Liver	30.6	0.0
Brain (whole)	31.4	0.0	Liver (fetal)	0.0	0.0
Brain (amygdala)	41.8	0.0	Liver ca. (hepatoblast) HepG2	0.0	0.0
Brain (cerebellum)	0.0	0.0	Lung	0.0	0.0
Brain (hippocampus)	37.1	0.0	Lung (fetal)	0.0	0.0
Brain (substantia nigra)	0.0	22.8	Lung ca. (small cell) LX-1	0.0	0.0
Brain (thalamus)	73.2	12.8	Lung ca. (small cell) NCI-H69	0.0	0.0
Cerebral Cortex	0.0	0.0	Lung ca. (s.cell var.) SHP-77	0.0	0.0
Spinal cord	0.0	0.0	Lung ca. (large cell) NCI-H460	0.0	0.0
glio/astro U87-MG	0.0	0.0	Lung ca. (non-sm. cell) A549	0.0	0.0
glio/astro U-118-MG	0.0	0.0	Lung ca. (non- s.cell) NCI-H23	33.7	0.0
astrocytoma SW1783	0.0	0.0	Lung ca. (non- s.cell) HOP-62	0.0	0.0
neuro*; met SK-N- AS	0.0	0.0	Lung ca. (non-s.cl) NCI-H522	0.0	0.0
astrocytoma SF-539	0.0	0.0	Lung ca. (squam.) SW 900	0.0	35.4
astrocytoma SNB-75	0.0	0.0	Lung ca. (squam.) NCI-H596	0.0	0.0
glioma SNB-19	0.0	0.0	Mammary gland	0.0	0.0
glioma U251	0.0	0.0	Breast ca.* (pl.ef)	0.0	100.0

			MCF-7		
glioma SF-295	0.0	0.0	Breast ca.* (pl.ef) MDA-MB-231	0.0	0.0
Heart (fetal)	0.0	24.7	Breast ca.* (pl.ef) T47D	0.0	0.0
Heart	0.0	0.0	Breast ca. BT-549	0.0	64.6
Skeletal muscle (fetal)	0.0	0.0	Breast ca. MDA-N	0.0	0.0
Skeletal muscle	0.0	32.3	Ovary	0.0	0.0
Bone marrow	0.0	29.5	Ovarian ca. OVCAR-3	37.9	0.0
Thymus	0.0	52.5	Ovarian ca. OVCAR-4	0.0	0.0
Spleen	0.0	0.0	Ovarian ca. OVCAR-5	0.0	0.0
Lymph node	100.0	94.6	Ovarian ca. OVCAR-8	0.0	0.0
Colorectal	46.7	82.9	Ovarian ca. IGROV-1	0.0	0.0
Stomach	0.0	0.0	Ovarian ca.* (ascites) SK-OV-3	0.0	39.8
Small intestine	0.0	0.0	Uterus	0.0	0.0
Colon ca. SW480	0.0	0.0	Placenta	41.2	12.6
Colon ca.* SW620(SW480 met)	0.0	0.0	Prostate	0.0	33.2
Colon ca. HT29	0.0	0.0	Prostate ca.* (bone met)PC-3	0.0	0.0
Colon ca. HCT-116	0.0	0.0	Testis	0.0	32.1
Colon ca. CaCo-2	0.0	0.0	Melanoma Hs688(A).T	0.0	0.0
Colon ca. tissue(ODO3866)	0.0	0.0	Melanoma* (met) Hs688(B).T	0.0	0.0
Colon ca. HCC-2998	0.0	0.0	Melanoma UACC- 62	0.0	0.0
Gastric ca.* (liver met) NCI-N87	0.0	0.0	Melanoma M14	0.0	0.0
Bladder	45.1	0.0	Melanoma LOX IMVI	0.0	0.0
Trachea	64.2	0.0	Melanoma* (met) SK-MEL-5	0.0	0.0
Kidney	40.9	0.0	Adipose	0.0	28.7

Table AWD. Panel 2D

Tissue Name	Rel. Exp.(%) Ag3046, Run 162559104	Tissue Name	Rel. Exp.(%) Ag3046, Run 162559104
Normal Colon	0.0	Kidney Margin 8120608	0.0
CC Well to Mod Diff (ODO3866)	0.0	Kidney Cancer 8120613	0.0
CC Margin (ODO3866)	0.0	Kidney Margin 8120614	0.0
CC Gr.2 rectosigmoid (ODO3868)	0.0	Kidney Cancer 9010320	0.0
CC Margin (ODO3868)	0.0	Kidney Margin 9010321	0.0
CC Mod Diff (ODO3920)	0.0	Normal Uterus	0.0
CC Margin (ODO3920)	0.0	Uterus Cancer 064011	0.0
CC Gr.2 ascend colon (ODO3921)	0.0	Normal Thyroid	0.0
CC Margin (ODO3921)	0.0	Thyroid Cancer 064010	0.0
CC from Partial Hepatectomy (ODO4309) Mets	0.0	Thyroid Cancer A302152	0.0
Liver Margin (ODO4309)	0.0	Thyroid Margin A302153	0.0
Colon mets to lung (OD04451- 01)	0.0	Normal Breast	0.0
Lung Margin (OD04451-02)	0.0	Breast Cancer (OD04566)	100.0
Normal Prostate 6546-1	0.1	Breast Cancer (OD04590-01)	0.0
Prostate Cancer (OD04410)	0.0	Breast Cancer Mets (OD04590- 03)	0.0
Prostate Margin (OD04410)	0.0	Breast Cancer Metastasis (OD04655-05)	0.0
Prostate Cancer (OD04720-01)	0.1	Breast Cancer 064006	0.0
Prostate Margin (OD04720- 02)	0.0	Breast Cancer 1024	0.0
Normal Lung 061010	0.0	Breast Cancer 9100266	0.0
Lung Met to Muscle (ODO4286)	0.0	Breast Margin 9100265	0.0
Muscle Margin (ODO4286)	0.0	Breast Cancer A209073	0.0
Lung Malignant Cancer (OD03126)	0.0	Breast Margin A209073	0.0
Lung Margin (OD03126)	0.0	Normal Liver	0.0
Lung Cancer (OD04404)	0.0	Liver Cancer 064003	0.0
Lung Margin (OD04404)	0.0	Liver Cancer 1025	0.0
Lung Cancer (OD04565)	0.0	Liver Cancer 1026	0.0
Lung Margin (OD04565)	0.0	Liver Cancer 6004-T	0.0
Lung Cancer (OD04237-01)	0.0	Liver Tissue 6004-N	0.0

Lung Margin (OD04237-02)	0.0	Liver Cancer 6005-T	0.0
Ocular Mel Met to Liver (ODO4310)	0.0	Liver Tissue 6005-N	0.0
Liver Margin (ODO4310)	0.0	Normal Bladder	0.0
Melanoma Mets to Lung (OD04321)	0.0	Bladder Cancer 1023	0.0
Lung Margin (OD04321)	0.0	Bladder Cancer A302173	0.0
Normal Kidney	0.0	Bladder Cancer (OD04718-01)	0.0
Kidney Ca, Nuclear grade 2 (OD04338)	0.0	Bladder Normal Adjacent (OD04718-03)	0.0
Kidney Margin (OD04338)	0.0	Normal Ovary	0.0
Kidney Ca Nuclear grade 1/2 (OD04339)	0.0	Ovarian Cancer 064008	0.0
Kidney Margin (OD04339)	0.0	Ovarian Cancer (OD04768-07)	0.0
Kidney Ca, Clear cell type (OD04340)	0.0	Ovary Margin (OD04768-08)	0.0
Kidney Margin (OD04340)	0.0	Normal Stomach	0.0
Kidney Ca, Nuclear grade 3 (OD04348)	0.0	Gastric Cancer 9060358	0.0
Kidney Margin (OD04348)	0.0	Stomach Margin 9060359	0.0
Kidney Cancer (OD04622-01)	0.0	Gastric Cancer 9060395	0.0
Kidney Margin (OD04622-03)	0.0	Stomach Margin 9060394	0.0
Kidney Cancer (OD04450-01)	0.0	Gastric Cancer 9060397	0.0
Kidney Margin (OD04450-03)	0.0	Stomach Margin 9060396	0.0
Kidney Cancer 8120607	0.0	Gastric Cancer 064005	0.0

Table AWE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4125, Run 172859315	Tissue Name	Rel. Exp.(%) Ag4125, Run 172859315
Secondary Th1 act	3.0	HUVEC IL-1beta	0.0
Secondary Th2 act	4.2	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	2.7	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	3.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	11.9	Microvascular Dermal EC TNFalpha + IL-1beta	7.4

Primary Th1 rest	7.2	Bronchial epithelium TNFalpha + IL1beta	7.5
Primary Th2 rest	3.1	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	3.3	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	5.3	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	1.8	NCI-H292 IL-4	5.8
LAK cells IL-2+ IL-18	3.6	NCI-H292 IL-9	3.5
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	1.3
NK Cells IL-2 rest	5.9	NCI-H292 IFN gamma	0.8
Two Way MLR 3 day	3.3	HPAEC none	0.0
Two Way MLR 5 day	3.2	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	2.3
PBMC PWM	3.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	1.6	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	1.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	2.1	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	3.3	Dermal fibroblast IFN gamma	0.0

Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	13.5	Colon	0.0
Macrophages rest	0.0	Lung	5.4
Macrophages LPS	0.0	Thymus	18.4
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

Table AWF. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag3046, Run 267908989	Rel. Exp.(%) Ag4125, Run 268390025	Tissue Name	Rel. Exp.(%) Ag3046, Run 267908989	Rel. Exp.(%) Ag4125, Run 268390025
Colon cancer 1	0.0	0.0	Bladder cancer NAT 2	0.0	0.0
Colon NAT 1	0.0	0.0	Bladder cancer NAT 3	0.0	3.6
Colon cancer 2	0.0	0.0	Bladder cancer NAT 4	0.0	7.6
Colon cancer NAT 2	0.0	0.0	Adenocarcinoma of the prostate 1	26.8	27.5
Colon cancer 3	12.1	14.1	Adenocarcinoma of the prostate 2	0.0	0.0
Colon cancer NAT 3	0.0	8.8	Adenocarcinoma of the prostate 3	0.0	18.4
Colon malignant cancer 4	11.7	23.8	Adenocarcinoma of the prostate 4	0.0	0.0
Colon normal adjacent tissue 4	3.5	0.0	Prostate cancer NAT 5	0.0	0.0
Lung cancer 1	0.0	14.2	Adenocarcinoma of the prostate 6	13.0	8.4
Lung NAT 1	7.3	0.0	Adenocarcinoma of the prostate 7	9.6	0.0
Lung cancer 2	8.7	0.0	Adenocarcinoma of the prostate 8	10.7	0.0
Lung NAT 2	0.0	0.0	Adenocarcinoma of the prostate 9	0.0	11.7
Squamous cell carcinoma 3	16.7	0.0	Prostate cancer NAT 10	0.0	0.0
Lung NAT 3	0.0	0.0	Kidney cancer 1	10.2	7.7
metastatic melanoma 1	8.8	14.8	Kidney NAT 1	0.0	0.0

Melanoma 2	9.1	0.0	Kidney cancer 2	100.0	100.0
Melanoma 3	16.2	8.2	Kidney NAT 2	0.0	7.7
metastatic melanoma 4	0.0	19.8	Kidney cancer 3	18.6	13.3
metastatic melanoma 5	58.2	33.7	Kidney NAT 3	0.0	8.1
Bladder cancer 1	0.0	0.0	Kidney cancer 4	17.8	0.0
Bladder cancer NAT 1	0.0	0.0	Kidney NAT 4	0.0	21.0
Bladder cancer 2	0.0	0.0			

CNS_neurodegeneration_v1.0 Summary: Ag3046 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Ag4125 Results from one experiment with this gene are not included. The amp plot indicates that there were experimental difficulties with this run (data not shown).

5 **General_screening_panel_v1.4 Summary:** Ag4125 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 1.3D Summary: Ag3046 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

10 **Panel 2D Summary:** Ag3046 Significant expression of this gene is seen exclusively in a breast cancer sample (CT = 25.2). Therefore, expression of this gene may be used to distinguish breast cancers from the other samples on this panel. Furthermore, therapeutic modulation of the activity of the GPCR encoded by this gene may be beneficial in the treatment of breast cancer.

15 **Panel 3D Summary:** Ag3046 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

20 **Panel 4.1D Summary:** Ag4125 This gene is only expressed at detectable levels in the kidney (CT = 32.6). The putative GPCR encoded for by this gene could allow cells within the kidney to respond to specific microenvironmental signals (For example, ref. 1). Therefore, antibody or small molecule therapies designed with the protein encoded for by this gene could modulate kidney function and be important in the treatment of inflammatory or autoimmune diseases that affect the kidney, including lupus and glomerulonephritis.

References:

1. Mark M.D., Wittemann S., Herlitze S. (2000) G protein modulation of recombinant P/Q-type calcium channels by regulators of G protein signalling proteins. J. Physiol. 528 Pt 1: 65-77.

1. Fast synaptic transmission is triggered by the activation of presynaptic Ca²⁺ channels which can be inhibited by Gbetagamma subunits via G protein-coupled receptors (GPCR). Regulators of G protein signalling (RGS) proteins are GTPase-accelerating proteins (GAPs), which are responsible for >100-fold increases in the GTPase activity of G proteins and might be involved in the regulation of presynaptic Ca²⁺ channels. In this study we investigated the effects of RGS2 on G protein modulation of recombinant P/Q-type channels expressed in a human embryonic kidney (HEK293) cell line using whole-cell recordings. 2. RGS2 markedly accelerates transmitter-mediated inhibition and recovery from inhibition of Ba²⁺ currents (IBa) through P/Q-type channels heterologously expressed with the muscarinic acetylcholine receptor M2 (mAChR M2). 3. Both RGS2 and RGS4 modulate the prepulse facilitation properties of P/Q-type Ca²⁺ channels. G protein reinhibition is accelerated, while release from inhibition is slowed. These kinetics depend on the availability of G protein alpha and betagamma subunits which is altered by RGS proteins. 4. RGS proteins unmask the Ca²⁺ channel beta subunit modulation of Ca²⁺ channel G protein inhibition. In the presence of RGS2, P/Q-type channels containing the beta2a and beta3 subunits reveal significantly altered kinetics of G protein modulation and increased facilitation compared to Ca²⁺ channels coexpressed with the beta1b or beta4 subunit.

PMID: 11018106

Panel 4D Summary: Ag3046 Expression of this gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

25 **General oncology screening panel_v_2.4 Summary:** Ag3046/Ag4125 Two experiments with same probe and primer set are in excellent agreement. Significant expression of this gene is seen exclusively in a kidney cancer sample (CT=34.6). Therefore, expression of this gene may be used to distinguish kidney cancers from the other samples on this panel. Furthermore, therapeutic modulation of the activity of the GPCR encoded by this gene may be beneficial in the treatment of kidney cancer.

AX. CG97658-01: Human Protein Tyrosine Phosphatase

Expression of gene CG97658-01 was assessed using the primer-probe set Ag4128, described in Table AXA. Results of the RTQ-PCR runs are shown in Tables AXB, AXC, AXD and AXE.

Table AXA. Probe Name Ag4128

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aagcagaagagcttcatgaaa-3'	22	784	386
Probe	TET-5'-cctttcctgctgcaggcggaatt-3'-TAMRA	23	816	387
Reverse	5'-aaagttcatggggatttcaaag-3'	22	839	388

5

Table AXB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4128, Run 214956370	Tissue Name	Rel. Exp.(%) Ag4128, Run 214956370
AD 1 Hippo	10.7	Control (Path) 3 Temporal Ctx	2.9
AD 2 Hippo	24.1	Control (Path) 4 Temporal Ctx	44.4
AD 3 Hippo	3.2	AD 1 Occipital Ctx	12.5
AD 4 Hippo	7.7	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	100.0	AD 3 Occipital Ctx	3.4
AD 6 Hippo	32.5	AD 4 Occipital Ctx	39.5
Control 2 Hippo	31.0	AD 5 Occipital Ctx	54.0
Control 4 Hippo	3.2	AD 6 Occipital Ctx	20.0
Control (Path) 3 Hippo	2.5	Control 1 Occipital Ctx	1.8
AD 1 Temporal Ctx	6.0	Control 2 Occipital Ctx	83.5
AD 2 Temporal Ctx	25.3	Control 3 Occipital Ctx	26.1
AD 3 Temporal Ctx	5.1	Control 4 Occipital Ctx	3.7
AD 4 Temporal Ctx	24.5	Control (Path) 1 Occipital Ctx	78.5
AD 5 Inf Temporal Ctx	92.0	Control (Path) 2 Occipital Ctx	23.8
AD 5 Sup Temporal Ctx	25.2	Control (Path) 3 Occipital Ctx	1.4
AD 6 Inf Temporal Ctx	33.7	Control (Path) 4 Occipital Ctx	23.0
AD 6 Sup Temporal Ctx	41.8	Control 1 Parietal Ctx	3.9
Control 1 Temporal Ctx	4.2	Control 2 Parietal Ctx	32.5
Control 2 Temporal Ctx	52.5	Control 3 Parietal Ctx	19.9
Control 3 Temporal Ctx	15.3	Control (Path) 1 Parietal Ctx	76.8
Control 3 Temporal Ctx	13.3	Control (Path) 2 Parietal Ctx	22.7
Control (Path) 1 Temporal Ctx	71.2	Control (Path) 3 Parietal Ctx	2.8
Control (Path) 2 Temporal Ctx	41.2	Control (Path) 4 Parietal Ctx	45.1

Table AXC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4128, Run 220364381	Tissue Name	Rel. Exp.(%) Ag4128, Run 220364381
Adipose	0.7	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.9
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.0
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.3	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	2.2	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.7	Colon ca. HCT-116	0.0
Prostate Pool	0.3	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.5
Uterus Pool	0.2	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.0	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.4	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	1.6
Ovarian ca. OVCAR-5	0.0	Small Intestine Pool	0.3
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.4
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.1
Ovary	6.0	Fetal Heart	2.2
Breast ca. MCF-7	0.0	Heart Pool	1.2
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	1.6
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.2
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.2
Breast ca. MDA-N	0.0	Spleen Pool	0.2
Breast Pool	0.6	Thymus Pool	0.3
Trachea	0.9	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.3	CNS cancer (glio/astro) U-118-MG	0.4
Fetal Lung	1.3	CNS cancer (neuro;met) SK-N-AS	0.1
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.0

Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	63.7
Lung ca. NCI-H526	0.0	Brain (cerebellum)	22.4
Lung ca. NCI-H23	0.1	Brain (fetal)	81.8
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	49.7
Lung ca. HOP-62	0.1	Cerebral Cortex Pool	77.9
Lung ca. NCI-H522	1.8	Brain (Substantia nigra) Pool	65.1
Liver	0.0	Brain (Thalamus) Pool	100.0
Fetal Liver	0.1	Brain (whole)	95.9
Liver ca. HepG2	0.0	Spinal Cord Pool	11.3
Kidney Pool	0.9	Adrenal Gland	0.6
Fetal Kidney	0.5	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.3
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.0	Pancreas Pool	0.8

Table AXD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4128, Run 172859562	Tissue Name	Rel. Exp.(%) Ag4128, Run 172859562
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	35.1
CD45RO CD4 lymphocyte	0.0	Coronary artery SMC	100.0

act		TNFalpha + IL-1beta	
CD8 lymphocyte act	0.0	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	9.9	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	10.4
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.0
Monocytes LPS	0.0	Colon	10.4

Macrophages rest	0.0	Lung	56.6
Macrophages LPS	0.0	Thymus	12.9
HUVEC none	0.0	Kidney	0.0
HUVEC starved	0.0		

Table AXE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4128, Run 268390026	Tissue Name	Rel. Exp.(%) Ag4128, Run 268390026
Colon cancer 1	4.1	Bladder cancer NAT 2	3.4
Colon NAT 1	4.3	Bladder cancer NAT 3	0.0
Colon cancer 2	3.8	Bladder cancer NAT 4	10.7
Colon cancer NAT 2	4.2	Adenocarcinoma of the prostate 1	4.2
Colon cancer 3	14.6	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	14.7	Adenocarcinoma of the prostate 3	2.1
Colon malignant cancer 4	9.3	Adenocarcinoma of the prostate 4	15.4
Colon normal adjacent tissue 4	1.7	Prostate cancer NAT 5	1.3
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	3.3
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	3.3
Lung cancer 2	43.8	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	1.6	Adenocarcinoma of the prostate 9	2.6
Squamous cell carcinoma 3	3.8	Prostate cancer NAT 10	1.7
Lung NAT 3	0.0	Kidney cancer 1	6.4
metastatic melanoma 1	42.9	Kidney NAT 1	9.7
Melanoma 2	3.7	Kidney cancer 2	29.1
Melanoma 3	2.1	Kidney NAT 2	9.7
metastatic melanoma 4	90.8	Kidney cancer 3	5.2
metastatic melanoma 5	100.0	Kidney NAT 3	2.0
Bladder cancer 1	0.0	Kidney cancer 4	18.3
Bladder cancer NAT 1	0.0	Kidney NAT 4	6.1
Bladder cancer 2	0.8		

CNS_neurodegeneration_v1.0 Summary: Ag4128 This panel confirms the expression of the CG97658-01 gene at low levels in the brains of an independent group of

individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

5 **General_screening_panel_v1.4 Summary:** Ag4128 Highest expression of the CG97658-01 gene is seen in thalamus (CT=25.7). High expression of this gene is detected in all regions of the central nervous system (CNS) examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, expression of this gene can be used to distinguish CNS samples from other
10 samples in this panel. Furthermore, therapeutic modulation of this gene may be useful in the treatment of central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Among tissues with metabolic or endocrine function, this gene is expressed at low to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal
15 muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Panel 4.1D Summary: Ag4128 Highest expression of the CG97658-01 gene is detected in TNFalpha + IL-1beta treated coronary artery SMC cells. Therefore, expression
20 of this gene can be used to distinguish this sample from other samples in this panel.

In addition, low expression of this gene is seen in lung. Therefore, therapeutic modulation of this gene product may be beneficial in the treatment of lung related disorders such as chronic obstructive pulmonary disease, asthma, allergy and emphysema.

General oncology screening panel_v_2.4 Summary: Ag4128 Highest
25 expression of the CG97658-01 gene is detected in metastatic melanoma (CT=31.5). Significant expression of this gene is associated with kidney cancer, melanoma and lung cancer. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

30 **AY. CG97842-01: Protein Kinase-Formin-Like Protein**

Expression of gene CG97842-01 was assessed using the primer-probe set Ag4130, described in Table AYA. Results of the RTQ-PCR runs are shown in Tables AYB, AYC and AYD.

Table AYA. Probe Name Ag4130

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-attcaacatgcccgtaataca-3'	22	2409	389
Probe	TET-5'-ccagaacttctacagtagtccaagcaca-3'-TAMRA	28	2433	390
Reverse	5'-agagtcattcttggtcactccaa-3'	22	2462	391

5

Table AYB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4130, Run 214959899	Tissue Name	Rel. Exp.(%) Ag4130, Run 214959899
AD 1 Hippo	16.4	Control (Path) 3 Temporal Ctx	7.2
AD 2 Hippo	34.6	Control (Path) 4 Temporal Ctx	50.0
AD 3 Hippo	9.6	AD 1 Occipital Ctx	21.5
AD 4 Hippo	16.3	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	90.8	AD 3 Occipital Ctx	8.7
AD 6 Hippo	80.1	AD 4 Occipital Ctx	20.2
Control 2 Hippo	33.0	AD 5 Occipital Ctx	37.1
Control 4 Hippo	13.8	AD 6 Occipital Ctx	55.9
Control (Path) 3 Hippo	17.0	Control 1 Occipital Ctx	4.4
AD 1 Temporal Ctx	20.7	Control 2 Occipital Ctx	36.3
AD 2 Temporal Ctx	42.3	Control 3 Occipital Ctx	25.0
AD 3 Temporal Ctx	10.2	Control 4 Occipital Ctx	10.7
AD 4 Temporal Ctx	40.1	Control (Path) 1 Occipital Ctx	85.3
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	24.8
AD 5 Sup Temporal Ctx	58.6	Control (Path) 3 Occipital Ctx	3.5
AD 6 Inf Temporal Ctx	69.3	Control (Path) 4 Occipital Ctx	29.9
AD 6 Sup Temporal Ctx	0.1	Control 1 Parietal Ctx	5.7
Control 1 Temporal Ctx	8.6	Control 2 Parietal Ctx	43.5
Control 2 Temporal Ctx	29.7	Control 3 Parietal Ctx	24.7
Control 3 Temporal Ctx	19.8	Control (Path) 1 Parietal Ctx	67.4
Control 3 Temporal Ctx	12.9	Control (Path) 2 Parietal Ctx	31.0
Control (Path) 1 Temporal Ctx	50.7	Control (Path) 3 Parietal Ctx	5.3
Control (Path) 2 Temporal Ctx	35.1	Control (Path) 4 Parietal Ctx	57.8

Table AYC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4130, Run 220424645	Tissue Name	Rel. Exp.(%) Ag4130, Run 220424645
Adipose	12.0	Renal ca. TK-10	34.2
Melanoma* Hs688(A).T	27.0	Bladder	28.5
Melanoma* Hs688(B).T	24.5	Gastric ca. (liver met.) NCI-N87	71.7
Melanoma* M14	41.5	Gastric ca. KATO III	47.6
Melanoma* LOXIMVI	25.2	Colon ca. SW-948	10.9
Melanoma* SK-MEL-5	33.9	Colon ca. SW480	65.5
Squamous cell carcinoma SCC-4	17.4	Colon ca.* (SW480 met) SW620	38.2
Testis Pool	12.9	Colon ca. HT29	27.7
Prostate ca.* (bone met) PC-3	95.9	Colon ca. HCT-116	43.8
Prostate Pool	8.7	Colon ca. CaCo-2	30.8
Placenta	3.5	Colon cancer tissue	37.4
Uterus Pool	4.2	Colon ca. SW1116	12.5
Ovarian ca. OVCAR-3	35.4	Colon ca. Colo-205	8.0
Ovarian ca. SK-OV-3	61.6	Colon ca. SW-48	7.6
Ovarian ca. OVCAR-4	27.0	Colon Pool	18.6
Ovarian ca. OVCAR-5	37.1	Small Intestine Pool	15.6
Ovarian ca. IGROV-1	18.4	Stomach Pool	10.7
Ovarian ca. OVCAR-8	14.9	Bone Marrow Pool	4.2
Ovary	11.3	Fetal Heart	18.7
Breast ca. MCF-7	59.9	Heart Pool	9.3
Breast ca. MDA-MB-231	46.7	Lymph Node Pool	17.0
Breast ca. BT 549	54.7	Fetal Skeletal Muscle	6.6
Breast ca. T47D	100.0	Skeletal Muscle Pool	13.8
Breast ca. MDA-N	18.3	Spleen Pool	12.2
Breast Pool	16.8	Thymus Pool	15.3
Trachea	9.5	CNS cancer (glio/astro) U87-MG	18.2
Lung	3.5	CNS cancer (glio/astro) U-118-MG	90.1
Fetal Lung	50.0	CNS cancer (neuro;met) SK-N-AS	57.8
Lung ca. NCI-N417	6.7	CNS cancer (astro) SF-539	21.2
Lung ca. LX-1	50.0	CNS cancer (astro) SNB-75	51.8

Lung ca. NCI-H146	13.3	CNS cancer (glio) SNB-19	16.4
Lung ca. SHP-77	30.4	CNS cancer (glio) SF-295	73.7
Lung ca. A549	37.4	Brain (Amygdala) Pool	6.1
Lung ca. NCI-H526	7.5	Brain (cerebellum)	8.2
Lung ca. NCI-H23	38.2	Brain (fetal)	23.5
Lung ca. NCI-H460	24.0	Brain (Hippocampus) Pool	9.1
Lung ca. HOP-62	13.0	Cerebral Cortex Pool	14.2
Lung ca. NCI-H522	46.3	Brain (Substantia nigra) Pool	6.2
Liver	1.1	Brain (Thalamus) Pool	15.6
Fetal Liver	28.1	Brain (whole)	5.3
Liver ca. HepG2	23.2	Spinal Cord Pool	8.1
Kidney Pool	21.9	Adrenal Gland	7.6
Fetal Kidney	28.3	Pituitary gland Pool	8.0
Renal ca. 786-0	43.8	Salivary Gland	1.9
Renal ca. A498	9.2	Thyroid (female)	4.0
Renal ca. ACHN	22.1	Pancreatic ca. CAPAN2	16.8
Renal ca. UO-31	32.3	Pancreas Pool	20.0

Table AYD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4130. Run 172859566	Tissue Name	Rel. Exp.(%) Ag4130, Run 172859566
Secondary Th1 act	39.8	HUVEC IL-1beta	43.8
Secondary Th2 act	46.3	HUVEC IFN gamma	48.3
Secondary Tr1 act	36.3	HUVEC TNF alpha + IFN gamma	34.2
Secondary Th1 rest	17.3	HUVEC TNF alpha + IL4	32.8
Secondary Th2 rest	15.8	HUVEC IL-11	30.8
Secondary Tr1 rest	19.1	Lung Microvascular EC none	57.0
Primary Th1 act	41.2	Lung Microvascular EC TNFalpha + IL-1beta	55.9
Primary Th2 act	34.2	Microvascular Dermal EC none	37.1
Primary Tr1 act	33.7	Microvascular Dermal EC TNFalpha + IL-1beta	37.1
Primary Th1 rest	28.5	Bronchial epithelium TNFalpha + IL1beta	100.0
Primary Th2 rest	14.5	Small airway epithelium none	12.4
Primary Tr1 rest	22.8	Small airway epithelium TNFalpha + IL-1beta	40.9

CD45RA CD4 lymphocyte act	35.8	Coronary artery SMC rest	24.3
CD45RO CD4 lymphocyte act	37.6	Coronary artery SMC TNFalpha + IL-1beta	23.8
CD8 lymphocyte act	37.1	Astrocytes rest	34.2
Secondary CD8 lymphocyte rest	26.4	Astrocytes TNFalpha + IL-1beta	22.8
Secondary CD8 lymphocyte act	16.2	KU-812 (Basophil) rest	36.6
CD4 lymphocyte none	12.2	KU-812 (Basophil) PMA/ionomycin	73.7
2ry Th1/Th2/Tr1_anti-CD95 CH11	31.9	CCD1106 (Keratinocytes) none	30.1
LAK cells rest	36.9	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	26.1
LAK cells IL-2	30.1	Liver cirrhosis	16.2
LAK cells IL-2+IL-12	21.8	NCI-H292 none	24.5
LAK cells IL-2+IFN gamma	33.9	NCI-H292 IL-4	46.3
LAK cells IL-2+ IL-18	29.5	NCI-H292 IL-9	48.0
LAK cells PMA/ionomycin	81.2	NCI-H292 IL-13	41.8
NK Cells IL-2 rest	36.1	NCI-H292 IFN gamma	39.2
Two Way MLR 3 day	35.4	HPAEC none	30.4
Two Way MLR 5 day	31.9	HPAEC TNF alpha + IL-1 beta	51.4
Two Way MLR 7 day	21.8	Lung fibroblast none	39.8
PBMC rest	14.6	Lung fibroblast TNF alpha + IL-1 beta	27.0
PBMC PWM	27.9	Lung fibroblast IL-4	27.5
PBMC PHA-L	23.7	Lung fibroblast IL-9	36.6
Ramos (B cell) none	30.4	Lung fibroblast IL-13	31.4
Ramos (B cell) ionomycin	35.4	Lung fibroblast IFN gamma	46.0
B lymphocytes PWM	23.3	Dermal fibroblast CCD1070 rest	45.7
B lymphocytes CD40L and IL-4	23.2	Dermal fibroblast CCD1070 TNF alpha	61.6
EOL-1 dbcAMP	29.5	Dermal fibroblast CCD1070 IL-1 beta	30.4
EOL-1 dbcAMP PMA/ionomycin	61.1	Dermal fibroblast IFN gamma	24.8
Dendritic cells none	57.0	Dermal fibroblast IL-4	36.9
Dendritic cells LPS	38.2	Dermal Fibroblasts rest	17.1
Dendritic cells anti-CD40	39.8	Neutrophils TNFa+LPS	11.3
Monocytes rest	38.4	Neutrophils rest	30.1
Monocytes LPS	72.2	Colon	10.0

Macrophages rest	36.9	Lung	17.9
Macrophages LPS	36.1	Thymus	47.3
HUVEC none	31.0	Kidney	44.8
HUVEC starved	38.4		

CNS_neurodegeneration_v1.0 Summary: Ag4130 This panel confirms the expression of the CG97842-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between
5 Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4130 Highest expression of the CG97842-01 gene is detected in a breast cancer cell line (CT=25.6). High expression of
10 this gene is seen a cluster of cancer cell lines including CNS, colon, gastric, renal, lung, breast, ovarian, prostate, melanoma, squamous cell carcinoma and pancreatic cancer cell lines. Therefore, therapeutic modulation of this gene product may be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high
15 to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, this gene is expressed at much higher levels in fetal (CT=27) when
20 compared to adult liver and lung (CT=30-32). This observation suggests that expression of this gene can be used to distinguish fetal from adult liver and lung, respectively. In addition, the relative overexpression of this gene in fetal tissue suggests that the protein product may enhance growth or development of lung and liver in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the
25 protein encoded by this gene could be useful in treatment of liver and lung related diseases.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in

central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4130 Highest expression of the CG97842-01 gene is detected in TNFalpha + IL1beta treated bronchial epithelium (CT=28.7). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin. and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

AZ. CG98021-01: Synaptotagmin III

Expression of gene CG98021-01 was assessed using the primer-probe set Ag4138, described in Table AZA. Results of the RTQ-PCR runs are shown in Tables AZB, AZC, AZD and AZE.

Table AZA. Probe Name Ag4138

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-tcattcattcaacaaacatcca-3'	22	117	392
Probe	TET-5'-agcaccaactacgccctacgtgct-3'-TAMRA	24	90	393
Reverse	5'-gaactccaggaccttggtctct-3'	22	45	394

Table AZB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4138, Run 214961618	Tissue Name	Rel. Exp.(%) Ag4138, Run 214961618
AD 1 Hippo	8.0	Control (Path) 3 Temporal Ctx	1.7

AD 2 Hippo	17.8	Control (Path) 4 Temporal Ctx	19.5
AD 3 Hippo	1.9	AD 1 Occipital Ctx	6.4
AD 4 Hippo	2.9	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	100.0	AD 3 Occipital Ctx	1.3
AD 6 Hippo	29.3	AD 4 Occipital Ctx	14.7
Control 2 Hippo	18.7	AD 5 Occipital Ctx	18.8
Control 4 Hippo	2.0	AD 6 Occipital Ctx	60.7
Control (Path) 3 Hippo	1.4	Control 1 Occipital Ctx	0.5
AD 1 Temporal Ctx	4.6	Control 2 Occipital Ctx	76.8
AD 2 Temporal Ctx	17.2	Control 3 Occipital Ctx	5.3
AD 3 Temporal Ctx	2.4	Control 4 Occipital Ctx	0.9
AD 4 Temporal Ctx	15.3	Control (Path) 1 Occipital Ctx	79.6
AD 5 Inf Temporal Ctx	81.8	Control (Path) 2 Occipital Ctx	8.6
AD 5 Sup Temporal Ctx	24.1	Control (Path) 3 Occipital Ctx	0.4
AD 6 Inf Temporal Ctx	25.0	Control (Path) 4 Occipital Ctx	10.8
AD 6 Sup Temporal Ctx	28.9	Control 1 Parietal Ctx	1.1
Control 1 Temporal Ctx	0.3	Control 2 Parietal Ctx	33.7
Control 2 Temporal Ctx	38.7	Control 3 Parietal Ctx	8.2
Control 3 Temporal Ctx	9.2	Control (Path) 1 Parietal Ctx	78.5
Control 4 Temporal Ctx	1.7	Control (Path) 2 Parietal Ctx	29.7
Control (Path) 1 Temporal Ctx	57.8	Control (Path) 3 Parietal Ctx	0.7
Control (Path) 2 Temporal Ctx	36.1	Control (Path) 4 Parietal Ctx	38.4

Table AZC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4138, Run 220967147	Tissue Name	Rel. Exp.(%) Ag4138, Run 220967147
Adipose	0.0	Renal ca. TK-10	0.8
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.1	Gastric ca. (liver met.) NCI-N87	0.1
Melanoma* M14	0.1	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	2.1	Colon ca. SW480	0.6
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.1
Testis Pool	0.8	Colon ca. HT29	0.0

Prostate ca.* (bone met) PC-3	0.5	Colon ca. HCT-116	6.1
Prostate Pool	0.4	Colon ca. CaCo-2	0.7
Placenta	0.7	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0
Ovarian ca. OVCAR-3	0.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	5.4	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.5	Colon Pool	0.2
Ovarian ca. OVCAR-5	1.2	Small Intestine Pool	0.0
Ovarian ca. IGROV-1	0.8	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.0	Bone Marrow Pool	0.0
Ovary	0.3	Fetal Heart	0.1
Breast ca. MCF-7	0.3	Heart Pool	0.0
Breast ca. MDA-MB-231	0.0	Lymph Node Pool	0.6
Breast ca. BT 549	0.0	Fetal Skeletal Muscle	0.0
Breast ca. T47D	1.1	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.1	Spleen Pool	0.1
Breast Pool	0.2	Thymus Pool	0.0
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.1
Fetal Lung	0.9	CNS cancer (neuro;met) SK-N-AS	1.1
Lung ca. NCI-N417	0.2	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	0.0
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.3
Lung ca. SHP-77	0.0	CNS cancer (glio) SF-295	0.7
Lung ca. A549	0.7	Brain (Amygdala) Pool	11.7
Lung ca. NCI-H526	0.0	Brain (cerebellum)	98.6
Lung ca. NCI-H23	3.7	Brain (fetal)	100.0
Lung ca. NCI-H460	0.5	Brain (Hippocampus) Pool	12.3
Lung ca. HOP-62	0.1	Cerebral Cortex Pool	24.5
Lung ca. NCI-H522	3.2	Brain (Substantia nigra) Pool	17.7
Liver	0.0	Brain (Thalamus) Pool	29.3
Fetal Liver	0.0	Brain (whole)	18.3
Liver ca. HepG2	0.0	Spinal Cord Pool	3.0
Kidney Pool	0.1	Adrenal Gland	1.0
Fetal Kidney	1.2	Pituitary gland Pool	1.9
Renal ca. 786-0	0.0	Salivary Gland	0.1

Renal ca. A498	1.1	Thyroid (female)	0.7
Renal ca. ACHN	0.2	Pancreatic ca. CAPAN2	0.0
Renal ca. UO-31	0.3	Pancreas Pool	0.2

Table AZD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4138, Run 173118876	Tissue Name	Rel. Exp.(%) Ag4138, Run 173118876
Secondary Th1 act	0.3	HUVEC IL-1beta	0.5
Secondary Th2 act	0.0	HUVEC IFN gamma	1.4
Secondary Tr1 act	0.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.7
Secondary Tr1 rest	0.0	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.8	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.3	Coronary artery SMC TNFalpha + IL-1beta	0.7
CD8 lymphocyte act	0.0	Astrocytes rest	0.4
Secondary CD8 lymphocyte rest	0.5	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.8
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	4.5
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.7	Liver cirrhosis	0.1
LAK cells IL-2+IL-12	0.0	NCI-H292 none	0.0

LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	1.0	Lung fibroblast none	0.9
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.5	Lung fibroblast IL-4	0.0
PBMC PHA-L	2.0	Lung fibroblast IL-9	0.9
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.4	Dermal fibroblast CCD1070 TNF alpha	0.4
EOL-1 dbcAMP	2.7	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.5	Dermal fibroblast IFN gamma	1.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.5
Dendritic cells LPS	1.0	Dermal Fibroblasts rest	2.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	0.0	Neutrophils rest	0.8
Monocytes LPS	0.0	Colon	0.8
Macrophages rest	0.0	Lung	0.9
Macrophages LPS	0.0	Thymus	13.4
HUVEC none	0.2	Kidney	100.0
HUVEC starved	0.5		

Table AZE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4138, Run 268390088	Tissue Name	Rel. Exp.(%) Ag4138, Run 268390088
Colon cancer 1	3.1	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	0.0
Colon cancer 2	11.8	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	7.6	Adenocarcinoma of the prostate 1	0.0

Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	14.5
Colon malignant cancer 4	32.1	Adenocarcinoma of the prostate 4	5.4
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	10.4
Lung cancer 1	6.0	Adenocarcinoma of the prostate 6	29.7
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	74.7	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	3.7	Adenocarcinoma of the prostate 9	6.5
Squamous cell carcinoma 3	21.8	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	6.2	Kidney NAT 1	4.2
Melanoma 2	0.0	Kidney cancer 2	24.3
Melanoma 3	0.0	Kidney NAT 2	100.0
metastatic melanoma 4	8.3	Kidney cancer 3	26.4
metastatic melanoma 5	11.0	Kidney NAT 3	7.5
Bladder cancer 1	0.0	Kidney cancer 4	11.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	8.5
Bladder cancer 2	4.3		

CNS_neurodegeneration_v1.0 Summary: Ag4138 This panel confirms the expression of the CG98021-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4138 Highest expression of the CG98021-01 gene is detected in cerebellum and fetal brain (CT=28). In addition high expression of this gene is detected exclusively in all the region of central nervous system examined. Therefore, expression of this gene can be used to distinguish CNS samples from other samples used in this panel. In addition, therapeutic modulation of this gene may be useful in the treatment of CNS disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Low expression of this gene is also seen in a few of colon cancer, renal cancer, lung cancer, breast, ovarian and melanoma cell lines. Therefore, therapeutic modulation of this gene may be beneficial in the treatment of these cancers.

Low expression of this gene is also detected in fetal lung. Interestingly, this gene is expressed at much higher levels in fetal (CT=34.8) when compared to adult lung (CT=40). This observation suggests that expression of this gene can be used to distinguish fetal from adult lung. In addition, the relative overexpression of this gene in fetal lung suggests that the protein product may enhance lung growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of lung related diseases.

Panel 4.1D Summary: Ag4138 Highest expression of the CG98021-01 gene is detected in kidney (CT=29.6). In addition low expression of this gene is also detected in thymus, eosinophils and PMA/ionomycin treated basophils. Therefore, expression of this gene can be used to distinguish these samples from other samples used in this panel. In addition, therapeutic modulation of this gene may be useful in the treatment of inflammation and autoimmune disease that affect kidney including lupus and glomerulonephritis.

General oncology screening panel_v_2.4 Summary: Ag4138 Highest expression of the CG98021-01 gene is detected in kidney (CT=34.5). In addition, low expression of this gene is also seen in lung cancer sample. Please see Panel 1.4 and 4.1D for a discussion of the potential utility of this gene.

20 **BA. CG98030-01: Tyrosine Phosphatase**

Expression of gene CG98030-01 was assessed using the primer-probe set Ag4139, described in Table BAA. Results of the RTQ-PCR runs are shown in Tables BAB and BAC.

Table BAA. Probe Name Ag4139

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - ccaattccatacaaaccaagag - 3'	22	548	395
Probe	TET-5' - actcagtttctaactcctctccgcaa - 3' - TAMRA	26	595	396
Reverse	5' - tgggatcacagcaagagaatat - 3'	22	622	397

25

Table BAB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4139, Run 214962531	Tissue Name	Rel. Exp.(%) Ag4139, Run 214962531

AD 1 Hippo	21.2	Control (Path) 3 Temporal Ctx	0.0
AD 2 Hippo	12.4	Control (Path) 4 Temporal Ctx	12.8
AD 3 Hippo	6.2	AD 1 Occipital Ctx	13.4
AD 4 Hippo	18.6	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	62.4	AD 3 Occipital Ctx	10.8
AD 6 Hippo	73.7	AD 4 Occipital Ctx	41.5
Control 2 Hippo	21.0	AD 5 Occipital Ctx	38.2
Control 4 Hippo	14.8	AD 6 Occipital Ctx	30.4
Control (Path) 3 Hippo	1.7	Control 1 Occipital Ctx	6.3
AD 1 Temporal Ctx	21.9	Control 2 Occipital Ctx	28.1
AD 2 Temporal Ctx	26.6	Control 3 Occipital Ctx	19.5
AD 3 Temporal Ctx	5.3	Control 4 Occipital Ctx	1.5
AD 4 Temporal Ctx	6.6	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	56.3	Control (Path) 2 Occipital Ctx	6.6
AD 5 Sup Temporal Ctx	84.7	Control (Path) 3 Occipital Ctx	2.5
AD 6 Inf Temporal Ctx	91.4	Control (Path) 4 Occipital Ctx	4.8
AD 6 Sup Temporal Ctx	44.8	Control 1 Parietal Ctx	3.7
Control 1 Temporal Ctx	1.1	Control 2 Parietal Ctx	37.4
Control 2 Temporal Ctx	55.9	Control 3 Parietal Ctx	4.5
Control 3 Temporal Ctx	27.5	Control (Path) 1 Parietal Ctx	62.9
Control 3 Temporal Ctx	3.1	Control (Path) 2 Parietal Ctx	24.8
Control (Path) 1 Temporal Ctx	31.9	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	28.7	Control (Path) 4 Parietal Ctx	75.8

Table BAC. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4139, Run 268392257	Tissue Name	Rel. Exp.(%) Ag4139, Run 268392257
Colon cancer 1	24.7	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	0.7	Bladder cancer NAT 3	1.1
Colon cancer 2	19.2	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.6	Adenocarcinoma of the prostate 1	2.0
Colon cancer 3	43.8	Adenocarcinoma of the prostate 2	1.1
Colon cancer NAT 3	1.0	Adenocarcinoma of the prostate 3	2.8
Colon malignant cancer 4	100.0	Adenocarcinoma of the prostate 4	9.1
Colon normal adjacent	0.0	Prostate cancer NAT 5	0.4

tissue 4			
Lung cancer 1	11.0	Adenocarcinoma of the prostate 6	0.8
Lung NAT 1	0.8	Adenocarcinoma of the prostate 7	0.8
Lung cancer 2	25.7	Adenocarcinoma of the prostate 8	0.6
Lung NAT 2	1.6	Adenocarcinoma of the prostate 9	2.2
Squamous cell carcinoma 3	18.2	Prostate cancer NAT 10	0.3
Lung NAT 3	0.2	Kidney cancer 1	3.4
metastatic melanoma 1	4.1	Kidney NAT 1	1.0
Melanoma 2	2.0	Kidney cancer 2	9.5
Melanoma 3	0.8	Kidney NAT 2	2.3
metastatic melanoma 4	4.5	Kidney cancer 3	1.7
metastatic melanoma 5	3.1	Kidney NAT 3	0.4
Bladder cancer 1	0.1	Kidney cancer 4	6.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	1.4
Bladder cancer 2	0.8		

CNS_neurodegeneration_v1.0 Summary: Ag4139 This panel confirms the expression of the CG98030-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

Panel 4.1D Summary: Ag4139 Expression of the CG98030-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4139 Highest expression of the CG98030-01 gene is detected in malignant colon cancer sample (CT=29). Interestingly, higher expression of this gene is associated with number of cancer samples examined including colon, lung, melanoma, metastatic melanoma, adenocarcinoma of the prostate, and kidney cancers. Therefore, expression of this gene may be used as diagnostic markers for these cancers. In addition, therapeutic modulation of this gene may be useful in the treatments of these cancers.

BB. CG98030-02: Tyrosine Phosphatase

Expression of full length physical clone CG98030-02 was assessed using the primer-probe set Ag6401, described in Table BBA. Results of the RTQ-PCR runs are shown in Table BBB.

Table BBA. Probe Name Ag6401

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-atgaatctgtagaaaaggagggaactaa-3'	27	1676	398
Probe	TET-5'-tctttctgccacatttctaccttct-3'-TAMRA	26	1705	399
Reverse	5'-ctccatctcggaattaagc-3'	20	1731	400

5

Table BBB. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag6401, Run 269239989	Tissue Name	Rel. Exp.(%) Ag6401, Run 269239989
Secondary Th1 act	17.2	HUVEC IL-1beta	37.6
Secondary Th2 act	17.2	HUVEC IFN gamma	33.9
Secondary Tr1 act	5.0	HUVEC TNF alpha + IFN gamma	16.5
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	23.8
Secondary Th2 rest	4.9	HUVEC IL-11	13.4
Secondary Tr1 rest	3.7	Lung Microvascular EC none	100.0
Primary Th1 act	58.2	Lung Microvascular EC TNFalpha + IL-1beta	32.8
Primary Th2 act	28.3	Microvascular Dermal EC none	14.1
Primary Tr1 act	35.4	Microvascular Dermal EC TNFalpha + IL-1beta	9.3
Primary Th1 rest	1.8	Bronchial epithelium TNFalpha + IL1beta	21.0
Primary Th2 rest	2.5	Small airway epithelium none	12.2
Primary Tr1 rest	3.2	Small airway epithelium TNFalpha + IL-1beta	11.9
CD45RA CD4 lymphocyte act	30.6	Coronary artery SMC rest	37.9
CD45RO CD4 lymphocyte act	54.3	Coronary artery SMC TNFalpha + IL-1beta	49.0
CD8 lymphocyte act	37.1	Astrocytes rest	19.9
Secondary CD8 lymphocyte rest	5.8	Astrocytes TNFalpha + IL-1beta	12.9
Secondary CD8 lymphocyte act	5.4	KU-812 (Basophil) rest	13.4
CD4 lymphocyte none	3.6	KU-812 (Basophil) PMA/ionomycin	5.3

2ry Th1/Th2/Tr1_anti-CD95 CH11	8.0	CCD1106 (Keratinocytes) none	97.9
LAK cells rest	21.8	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	41.5
LAK cells IL-2	5.6	Liver cirrhosis	17.3
LAK cells IL-2+IL-12	1.6	NCI-H292 none	13.3
LAK cells IL-2+IFN gamma	3.4	NCI-H292 IL-4	33.4
LAK cells IL-2+ IL-18	6.0	NCI-H292 IL-9	68.3
LAK cells PMA/ionomycin	6.9	NCI-H292 IL-13	24.7
NK Cells IL-2 rest	16.3	NCI-H292 IFN gamma	25.0
Two Way MLR 3 day	5.1	HPAEC none	20.2
Two Way MLR 5 day	2.5	HPAEC TNF alpha + IL-1 beta	64.6
Two Way MLR 7 day	2.3	Lung fibroblast none	61.1
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	16.5
PBMC PWM	1.9	Lung fibroblast IL-4	17.0
PBMC PHA-L	17.7	Lung fibroblast IL-9	14.1
Ramos (B cell) none	6.8	Lung fibroblast IL-13	9.9
Ramos (B cell) ionomycin	9.5	Lung fibroblast IFN gamma	15.2
B lymphocytes PWM	8.0	Dermal fibroblast CCD1070 rest	78.5
B lymphocytes CD40L and IL-4	15.2	Dermal fibroblast CCD1070 TNF alpha	63.3
EOL-1 dbcAMP	6.2	Dermal fibroblast CCD1070 IL-1 beta	47.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	9.3
Dendritic cells none	10.2	Dermal fibroblast IL-4	47.6
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	39.8
Dendritic cells anti-CD40	10.2	Neutrophils TNFa+LPS	1.8
Monocytes rest	1.3	Neutrophils rest	4.1
Monocytes LPS	5.9	Colon	3.3
Macrophages rest	2.5	Lung	6.7
Macrophages LPS	0.0	Thymus	4.0
HUVEC none	16.8	Kidney	38.7
HUVEC starved	54.3		

Panel 4.1D Summary: Ag6401 Highest expression of the CG98030-02 is detected in lung microvascular endothelial cells and keratinocytes (CT=32.8). In addition, low levels of expression of this gene is also seen in dermal fibroblast, lung fibroblast, TNF

alpha + IL-1 beta treated HPAEC, HUVEC, resting LAK cells, activated primary Th1, Th2, and Tr1 cells, activated CD4 lymphocytes and kidney. Therefore, therapeutic modulation of this tyrosine phosphatase encoded by this gene may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

BC. CG98061-01 and CG98061-02: Novel Protein containing Histidine acid phosphatase domain

10 Expression of gene CG98061-01 and variant CG98061-02 was assessed using the primer-probe set Ag4141, described in Table BCA. Results of the RTQ-PCR runs are shown in Tables BCB, BCC and BCD.

Table BCA. Probe Name Ag4141

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-gtgcttggcaactgatgttc-3'	20	2241	401
Probe	TET-5'-tgacacactggagtctaataccaaca-3'-TAMRA	27	2265	402
Reverse	5'-gcaccagagtctgttaattcca-3'	22	2303	403

Table BCB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4141, Run 215281611	Tissue Name	Rel. Exp.(%) Ag4141, Run 215281611
AD 1 Hippo	15.9	Control (Path) 3 Temporal Ctx	4.4
AD 2 Hippo	20.7	Control (Path) 4 Temporal Ctx	13.1
AD 3 Hippo	6.6	AD 1 Occipital Ctx	13.1
AD 4 Hippo	2.8	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	58.6	AD 3 Occipital Ctx	6.7
AD 6 Hippo	100.0	AD 4 Occipital Ctx	12.4
Control 2 Hippo	17.7	AD 5 Occipital Ctx	18.0
Control 4 Hippo	13.9	AD 6 Occipital Ctx	21.9
Control (Path) 3 Hippo	7.1	Control 1 Occipital Ctx	6.0
AD 1 Temporal Ctx	13.5	Control 2 Occipital Ctx	37.4
AD 2 Temporal Ctx	23.3	Control 3 Occipital Ctx	11.8
AD 3 Temporal Ctx	1.6	Control 4 Occipital Ctx	8.0
AD 4 Temporal Ctx	12.9	Control (Path) 1 Occipital Ctx	48.6

AD 5 Inf Temporal Ctx	62.4	Control (Path) 2 Occipital Ctx	6.3
AD 5 Sup Temporal Ctx	48.3	Control (Path) 3 Occipital Ctx	5.3
AD 6 Inf Temporal Ctx	83.5	Control (Path) 4 Occipital Ctx	7.5
AD 6 Sup Temporal Ctx	70.7	Control 1 Parietal Ctx	7.0
Control 1 Temporal Ctx	4.5	Control 2 Parietal Ctx	40.9
Control 2 Temporal Ctx	18.7	Control 3 Parietal Ctx	11.5
Control 3 Temporal Ctx	10.5	Control (Path) 1 Parietal Ctx	37.9
Control 3 Temporal Ctx	7.5	Control (Path) 2 Parietal Ctx	18.3
Control (Path) 1 Temporal Ctx	30.8	Control (Path) 3 Parietal Ctx	5.4
Control (Path) 2 Temporal Ctx	22.1	Control (Path) 4 Parietal Ctx	29.3

Table BCC. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4141, Run 173118944	Tissue Name	Rel. Exp.(%) Ag4141, Run 173118944
Secondary Th1 act	82.4	HUVEC IL-1beta	19.8
Secondary Th2 act	100.0	HUVEC IFN gamma	17.6
Secondary Tr1 act	62.0	HUVEC TNF alpha + IFN gamma	15.4
Secondary Th1 rest	6.9	HUVEC TNF alpha + IL4	25.3
Secondary Th2 rest	11.8	HUVEC IL-11	8.3
Secondary Tr1 rest	11.2	Lung Microvascular EC none	24.5
Primary Th1 act	61.1	Lung Microvascular EC TNFalpha + IL-1beta	17.2
Primary Th2 act	88.3	Microvascular Dermal EC none	14.1
Primary Tr1 act	88.3	Microvascular Dermal EC TNFalpha + IL-1beta	11.8
Primary Th1 rest	11.2	Bronchial epithelium TNFalpha + IL1beta	12.6
Primary Th2 rest	5.8	Small airway epithelium none	5.8
Primary Tr1 rest	15.4	Small airway epithelium TNFalpha + IL-1beta	10.9
CD45RA CD4 lymphocyte act	55.9	Coronary artery SMC rest	9.4
CD45RO CD4 lymphocyte act	91.4	Coronary artery SMC TNFalpha + IL-1beta	8.4
CD8 lymphocyte act	60.7	Astrocytes rest	10.2
Secondary CD8 lymphocyte rest	56.6	Astrocytes TNFalpha + IL-1beta	8.2
Secondary CD8 lymphocyte act	21.8	KU-812 (Basophil) rest	35.4
CD4 lymphocyte none	4.3	KU-812 (Basophil)	66.0

		PMA/ionomycin	
2ry Th1/Th2/Tr1_anti-CD95 CH11	14.6	CCD1106 (Keratinocytes) none	33.4
LAK cells rest	19.2	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	31.2
LAK cells IL-2	28.9	Liver cirrhosis	2.7
LAK cells IL-2+IL-12	33.7	NCI-H292 none	26.6
LAK cells IL-2+IFN gamma	18.3	NCI-H292 IL-4	63.7
LAK cells IL-2+ IL-18	16.6	NCI-H292 IL-9	69.7
LAK cells PMA/ionomycin	20.4	NCI-H292 IL-13	69.3
NK Cells IL-2 rest	40.1	NCI-H292 IFN gamma	67.8
Two Way MLR 3 day	24.3	HPAEC none	11.7
Two Way MLR 5 day	29.7	HPAEC TNF alpha + IL-1 beta	24.1
Two Way MLR 7 day	23.8	Lung fibroblast none	10.9
PBMC rest	3.4	Lung fibroblast TNF alpha + IL-1 beta	12.5
PBMC PWM	30.4	Lung fibroblast IL-4	24.1
PBMC PHA-L	31.6	Lung fibroblast IL-9	22.8
Ramos (B cell) none	26.1	Lung fibroblast IL-13	24.3
Ramos (B cell) ionomycin	30.1	Lung fibroblast IFN gamma	39.2
B lymphocytes PWM	43.2	Dermal fibroblast CCD1070 rest	42.3
B lymphocytes CD40L and IL- 4	25.0	Dermal fibroblast CCD1070 TNF alpha	54.3
EOL-1 dbcAMP	51.8	Dermal fibroblast CCD1070 IL-1 beta	26.1
EOL-1 dbcAMP PMA/ionomycin	20.2	Dermal fibroblast IFN gamma	29.1
Dendritic cells none	14.4	Dermal fibroblast IL-4	36.9
Dendritic cells LPS	12.1	Dermal Fibroblasts rest	15.0
Dendritic cells anti-CD40	15.9	Neutrophils TNFa+LPS	2.0
Monocytes rest	7.1	Neutrophils rest	5.6
Monocytes LPS	28.9	Colon	5.8
Macrophages rest	10.9	Lung	11.8
Macrophages LPS	9.1	Thymus	24.5
HUVEC none	14.6	Kidney	14.5
HUVEC starved	21.9		

Table BCD. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4141, Run 268392261	Tissue Name	Rel. Exp.(%) Ag4141,
-------------	--	-------------	----------------------------

			Run 268392261
Colon cancer 1	0.6	Bladder cancer NAT 2	0.0
Colon cancer NAT 1	0.0	Bladder cancer NAT 3	0.8
Colon cancer 2	0.0	Bladder cancer NAT 4	1.3
Colon cancer NAT 2	0.6	Adenocarcinoma of the prostate 1	7.5
Colon cancer 3	1.3	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	1.3	Adenocarcinoma of the prostate 3	10.4
Colon malignant cancer 4	8.7	Adenocarcinoma of the prostate 4	3.4
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	5.4
Lung cancer 1	8.4	Adenocarcinoma of the prostate 6	5.2
Lung NAT 1	3.4	Adenocarcinoma of the prostate 7	1.1
Lung cancer 2	23.7	Adenocarcinoma of the prostate 8	1.4
Lung NAT 2	1.8	Adenocarcinoma of the prostate 9	17.4
Squamous cell carcinoma 3	2.3	Prostate cancer NAT 10	0.4
Lung NAT 3	3.2	Kidney cancer 1	20.6
metastatic melanoma 1	29.5	Kidney NAT 1	4.6
Melanoma 2	0.5	Kidney cancer 2	100.0
Melanoma 3	0.0	Kidney NAT 2	7.8
metastatic melanoma 4	6.4	Kidney cancer 3	31.6
metastatic melanoma 5	6.4	Kidney NAT 3	4.5
Bladder cancer 1	0.9	Kidney cancer 4	3.1
Bladder cancer NAT 1	0.0	Kidney NAT 4	1.5
Bladder cancer 2	0.9		

CNS_neurodegeneration_v1.0 Summary: Ag4141 This panel confirms the expression of the CG98061-01 gene at low levels in the brain in an independent group of individuals. This gene is found to be slightly upregulated in the temporal cortex of Alzheimer's disease patients. Therefore, therapeutic modulation of the expression or
5 function of this gene may decrease neuronal death and be of use in the treatment of this disease.

General_screening_panel_v1.4 Summary: Ag4141 Results from one experiment with the CG98061-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

10 **Panel 4.1D Summary:** Ag4141 Highest expression of the CG98061-01 gene is detected in activated secondary Th2 cells (CT=29.6). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in

health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4141 Highest expression of the CG98061-01 gene is detected in kidney cancer sample (CT=31.8). In addition, significant expression of this gene is also seen in lung cancer, metastatic melanoma, kidney cancer and adenocarcinoma of the prostate. Therefore, therapeutic modulation of this gene may be beneficial in the treatment of these cancers.

BD. CG98131-01: MDJ6

Expression of gene CG98131-01 was assessed using the primer-probe set Ag4144, described in Table BDA. Results of the RTQ-PCR runs are shown in Tables BDB, BDC and BDD.

Table BDA. Probe Name Ag4144

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-cggaggacatcaagaaagc-3'	19	137	404
Probe	TET-5'-ctaccgcaagctggcccttcgtt-3'-TAMRA	23	156	405
Reverse	5'-tcctccttattgtcagggttct-3'	22	191	406

Table BDB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4144, Run 215294700	Tissue Name	Rel. Exp.(%) Ag4144, Run 215294700
AD 1 Hippo	0.0	Control (Path) 3 Temporal Ctx	0.0
AD 2 Hippo	0.0	Control (Path) 4 Temporal Ctx	0.0

AD 3 Hippo	0.0	AD 1 Occipital Ctx	0.0
AD 4 Hippo	0.0	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	0.0	AD 3 Occipital Ctx	0.0
AD 6 Hippo	0.0	AD 4 Occipital Ctx	0.0
Control 2 Hippo	0.0	AD 5 Occipital Ctx	0.0
Control 4 Hippo	0.0	AD 6 Occipital Ctx	0.0
Control (Path) 3 Hippo	24.8	Control 1 Occipital Ctx	0.0
AD 1 Temporal Ctx	26.6	Control 2 Occipital Ctx	13.2
AD 2 Temporal Ctx	0.0	Control 3 Occipital Ctx	0.0
AD 3 Temporal Ctx	0.0	Control 4 Occipital Ctx	0.0
AD 4 Temporal Ctx	21.6	Control (Path) 1 Occipital Ctx	19.5
AD 5 Inf Temporal Ctx	0.0	Control (Path) 2 Occipital Ctx	0.0
AD 5 Sup Temporal Ctx	0.0	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	0.0	Control (Path) 4 Occipital Ctx	0.0
AD 6 Sup Temporal Ctx	100.0	Control 1 Parietal Ctx	0.0
Control 1 Temporal Ctx	0.0	Control 2 Parietal Ctx	0.0
Control 2 Temporal Ctx	47.3	Control 3 Parietal Ctx	0.0
Control 3 Temporal Ctx	0.0	Control (Path) 1 Parietal Ctx	47.0
Control 4 Temporal Ctx	0.0	Control (Path) 2 Parietal Ctx	0.0
Control (Path) 1 Temporal Ctx	0.0	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	0.0	Control (Path) 4 Parietal Ctx	0.0

Table BDC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4144, Run 221000453	Tissue Name	Rel. Exp.(%) Ag4144, Run 221000453
Adipose	0.0	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.0	Bladder	0.0
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	1.1
Melanoma* M14	0.0	Gastric ca. KATO III	0.6
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.0
Melanoma* SK-MEL-5	0.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	100.0	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.0	Colon ca. HCT-116	0.9
Prostate Pool	0.0	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.0
Uterus Pool	0.0	Colon ca. SW1116	0.0

Ovarian ca. OVCAR-3	2.7	Colon ca. Colo-205	0.5
Ovarian ca. SK-OV-3	0.6	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.6
Ovarian ca. OVCAR-5	1.0	Small Intestine Pool	1.3
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.0
Ovarian ca. OVCAR-8	0.6	Bone Marrow Pool	0.5
Ovary	0.0	Fetal Heart	0.0
Breast ca. MCF-7	0.0	Heart Pool	0.0
Breast ca. MDA-MB-231	0.2	Lymph Node Pool	0.5
Breast ca. BT 549	0.8	Fetal Skeletal Muscle	0.0
Breast ca. T47D	0.0	Skeletal Muscle Pool	0.0
Breast ca. MDA-N	0.0	Spleen Pool	0.0
Breast Pool	0.2	Thymus Pool	2.0
Trachea	0.0	CNS cancer (glio/astro) U87-MG	0.0
Lung	0.0	CNS cancer (glio/astro) U-118-MG	0.0
Fetal Lung	0.0	CNS cancer (neuro;met) SK-N-AS	2.0
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	1.1
Lung ca. NCI-H146	0.6	CNS cancer (glio) SNB-19	0.0
Lung ca. SHP-77	0.6	CNS cancer (glio) SF-295	0.0
Lung ca. A549	0.0	Brain (Amygdala) Pool	0.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.5
Lung ca. NCI-H23	0.0	Brain (fetal)	1.5
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	0.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	1.5
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	0.5
Liver	0.0	Brain (Thalamus) Pool	0.0
Fetal Liver	0.0	Brain (whole)	0.4
Liver ca. HepG2	0.0	Spinal Cord Pool	0.0
Kidney Pool	2.6	Adrenal Gland	0.0
Fetal Kidney	0.0	Pituitary gland Pool	0.0
Renal ca. 786-0	0.0	Salivary Gland	0.0
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	2.5
Renal ca. UO-31	0.0	Pancreas Pool	0.7

Table BDD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4144, Run 173120070	Tissue Name	Rel. Exp.(%) Ag4144, Run 173120070
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	0.0	HUVEC IFN gamma	0.0
Secondary Tr1 act	3.4	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	2.2	Lung Microvascular EC none	0.0
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	3.7	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	0.0
Primary Th2 rest	0.0	Small airway epithelium none	0.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	3.1	Astrocytes rest	1.6
Secondary CD8 lymphocyte rest	3.8	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	2.7	KU-812 (Basophil) PMA/ionomycin	10.1
2ry Th1/Th2/Tr1_anti-CD95 CH11	0.0	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	0.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	7.8
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	6.9	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	12.2	NCI-H292 IFN gamma	0.0
Two Way MLR 3 day	0.0	HPAEC none	0.0

Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	3.6
PBMC rest	9.1	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	0.0
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	0.0
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	0.0	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	0.0
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	7.2	Neutrophils rest	0.0
Monocytes LPS	6.4	Colon	0.0
Macrophages rest	0.0	Lung	0.0
Macrophages LPS	0.0	Thymus	10.7
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4144 Highest expression of the CG98131-01 gene is detected in superior temporal cortex of an Alzheimer's disease patient (CT=34.1). Therefore, therapeutic modulation of the expression or function of this gene may decrease neuronal death and be of use in the treatment of this disease.

5 **General_screening_panel_v1.4 Summary:** Ag4144 Highest expression of the CG98131-01 gene is detected in testis (CT=29). Therefore, expression of this gene may be used to distinguish testis from other samples used in this panel and therapeutic modulation of this gene product may be useful in the treatment of disorders associated with testis such as fertility and hypogonadism.

10 In addition, low expression of this gene is also seen in a CNS cancer, pancreatic cancer and an ovarian cancer cell lines. Therefore, therapeutic modulation of this gene product may be beneficial in the treatment of these cancers.

Panel 4.1D Summary: Ag4144 Highest expression of the CG98131-01 gene is detected in exclusively in kidney (CT=32). Therefore, expression of this gene may be used to distinguish kidney from other samples used in this panel. In addition, therapeutic modulation of this gene product may be beneficial in the treatment of autoimmune and inflammatory diseases that affect kidney including lupus and glomerulonephritis.

BE. CG98164-01 and CG98164-02: LRR and Kinase Domain Protein

Expression of gene CG98164-01 and full length physical clone CG98164-02 was assessed using the primer-probe sets Ag4145 and Ag4145, described in Tables BEA and BEB. Results of the RTQ-PCR runs are shown in Tables BEC, BED, BEE and BEF. Please note that CG98164-02 represents a full-length physical clone of the CG98164-01 gene, validating the prediction of the gene sequence.

Table BEA. Probe Name Ag4145

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - agtgacagtgagaccgaagaga - 3'	22	2340	407
Probe	TET- 5' - cccggaaagcactacctatacaatca - 3' - TAMRA	26	2362	408
Reverse	5' - agtggtgtgtgtgtgttggtgaga - 3'	22	2406	409

Table BEB. Probe Name Ag4145

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5' - agtgacagtgagaccgaagaga - 3'	22	2340	410
Probe	TET- 5' - cccggaaagcactacctatacaatca - 3' - TAMRA	26	2362	411
Reverse	5' - agtggtgtgtgtgtgttggtgaga - 3'	22	2406	412

Table BEC. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4145, Run 215300524	Tissue Name	Rel. Exp.(%) Ag4145, Run 215300524
AD 1 Hippo	25.2	Control (Path) 3 Temporal Ctx	19.6
AD 2 Hippo	33.9	Control (Path) 4 Temporal Ctx	72.7
AD 3 Hippo	26.4	AD 1 Occipital Ctx	23.7
AD 4 Hippo	62.4	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	28.1	AD 3 Occipital Ctx	6.8
AD 6 Hippo	65.1	AD 4 Occipital Ctx	55.1
Control 2 Hippo	52.5	AD 5 Occipital Ctx	17.3

Control 4 Hippo	93.3	AD 6 Occipital Ctx	8.9
Control (Path) 3 Hippo	28.5	Control 1 Occipital Ctx	7.7
AD 1 Temporal Ctx	8.2	Control 2 Occipital Ctx	29.5
AD 2 Temporal Ctx	16.2	Control 3 Occipital Ctx	19.1
AD 3 Temporal Ctx	24.8	Control 4 Occipital Ctx	54.7
AD 4 Temporal Ctx	35.4	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	57.0	Control (Path) 2 Occipital Ctx	27.4
AD 5 Sup Temporal Ctx	66.9	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	34.2	Control (Path) 4 Occipital Ctx	15.6
AD 6 Sup Temporal Ctx	65.1	Control 1 Parietal Ctx	16.6
Control 1 Temporal Ctx	16.4	Control 2 Parietal Ctx	69.3
Control 2 Temporal Ctx	30.6	Control 3 Parietal Ctx	10.0
Control 3 Temporal Ctx	17.0	Control (Path) 1 Parietal Ctx	83.5
Control 3 Temporal Ctx	19.3	Control (Path) 2 Parietal Ctx	28.9
Control (Path) 1 Temporal Ctx	90.8	Control (Path) 3 Parietal Ctx	29.5
Control (Path) 2 Temporal Ctx	60.3	Control (Path) 4 Parietal Ctx	31.9

Table BED. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4145, Run 221000454	Tissue Name	Rel. Exp.(%) Ag4145, Run 221000454
Adipose	0.5	Renal ca. TK-10	0.4
Melanoma* Hs688(A).T	0.1	Bladder	8.1
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	3.4
Melanoma* M14	6.2	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	4.0	Colon ca. SW-948	0.2
Melanoma* SK-MEL-5	100.0	Colon ca. SW480	0.2
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.3
Testis Pool	23.2	Colon ca. HT29	0.3
Prostate ca.* (bone met) PC-3	0.8	Colon ca. HCT-116	5.6
Prostate Pool	1.5	Colon ca. CaCo-2	0.0
Placenta	0.0	Colon cancer tissue	0.4
Uterus Pool	0.7	Colon ca. SW1116	2.7
Ovarian ca. OVCAR-3	1.6	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	1.4	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.0	Colon Pool	0.9
Ovarian ca. OVCAR-5	1.3	Small Intestine Pool	2.3

Ovarian ca. IGROV-1	1.6	Stomach Pool	1.2
Ovarian ca. OVCAR-8	4.7	Bone Marrow Pool	0.6
Ovary	0.5	Fetal Heart	1.6
Breast ca. MCF-7	2.6	Heart Pool	0.4
Breast ca. MDA-MB-231	2.8	Lymph Node Pool	2.1
Breast ca. BT 549	5.8	Fetal Skeletal Muscle	1.2
Breast ca. T47D	0.9	Skeletal Muscle Pool	0.7
Breast ca. MDA-N	5.5	Spleen Pool	2.5
Breast Pool	1.7	Thymus Pool	2.2
Trachea	11.3	CNS cancer (glio/astro) U87-MG	0.7
Lung	0.0	CNS cancer (glio/astro) U-118-MG	1.3
Fetal Lung	17.6	CNS cancer (neuro:met) SK-N-AS	1.2
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.4	CNS cancer (astro) SNB-75	6.0
Lung ca. NCI-H146	0.2	CNS cancer (glio) SNB-19	1.0
Lung ca. SHP-77	0.8	CNS cancer (glio) SF-295	0.7
Lung ca. A549	1.5	Brain (Amygdala) Pool	1.0
Lung ca. NCI-H526	0.0	Brain (cerebellum)	1.3
Lung ca. NCI-H23	7.0	Brain (fetal)	3.7
Lung ca. NCI-H460	0.0	Brain (Hippocampus) Pool	1.0
Lung ca. HOP-62	0.0	Cerebral Cortex Pool	1.5
Lung ca. NCI-H522	0.5	Brain (Substantia nigra) Pool	0.5
Liver	0.0	Brain (Thalamus) Pool	3.1
Fetal Liver	1.0	Brain (whole)	1.4
Liver ca. HepG2	0.3	Spinal Cord Pool	1.7
Kidney Pool	2.3	Adrenal Gland	0.6
Fetal Kidney	7.0	Pituitary gland Pool	3.5
Renal ca. 786-0	0.3	Salivary Gland	0.8
Renal ca. A498	3.4	Thyroid (female)	1.0
Renal ca. ACHN	0.0	Pancreatic ca. CAPAN2	2.3
Renal ca. UO-31	0.9	Pancreas Pool	1.7

Table BEE. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4145, Run 197470572	Tissue Name	Rel. Exp.(%) Ag4145, Run 197470572
Secondary Th1 act	0.0	HUVEC IL-1beta	7.5
Secondary Th2 act	7.3	HUVEC IFN gamma	0.0

Secondary Tr1 act	23.0	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	0.0	Lung Microvascular EC none	11.3
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	8.2
Primary Tr1 act	7.9	Microvascular Dermal EC TNFalpha + IL-1beta	15.0
Primary Th1 rest	11.7	Bronchial epithelium TNFalpha + IL1beta	30.6
Primary Th2 rest	0.0	Small airway epithelium none	6.3
Primary Tr1 rest	7.7	Small airway epithelium TNFalpha + IL-1beta	15.5
CD45RA CD4 lymphocyte act	31.4	Coronary artery SMC rest	14.6
CD45RO CD4 lymphocyte act	40.1	Coronary artery SMC TNFalpha + IL-1beta	7.0
CD8 lymphocyte act	14.6	Astrocytes rest	0.0
Secondary CD8 lymphocyte rest	16.5	Astrocytes TNFalpha + IL-1beta	0.0
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	25.7
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	40.6
2ry Th1/Th2/Tr1_anti-CD95 CH11	8.6	CCD1106 (Keratinocytes) none	0.0
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	0.0
LAK cells IL-2	10.4	Liver cirrhosis	41.2
LAK cells IL-2+IL-12	11.6	NCI-H292 none	3.3
LAK cells IL-2+IFN gamma	13.8	NCI-H292 IL-4	0.0
LAK cells IL-2+ IL-18	16.0	NCI-H292 IL-9	0.0
LAK cells PMA/ionomycin	6.2	NCI-H292 IL-13	0.0
NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	5.5
Two Way MLR 3 day	0.0	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	6.7
Two Way MLR 7 day	0.0	Lung fibroblast none	0.0
PBMC rest	0.0	Lung fibroblast TNF alpha + IL-1 beta	7.1
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	8.4	Lung fibroblast IL-9	0.0
Ramos (B cell) none	94.6	Lung fibroblast IL-13	0.0

Ramos (B cell) ionomycin	100.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	13.4
B lymphocytes CD40L and IL-4	14.4	Dermal fibroblast CCD1070 TNF alpha	0.0
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	28.3	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	45.7	Dermal fibroblast IL-4	8.3
Dendritic cells LPS	16.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	23.5	Neutrophils TNFa+LPS	0.0
Monocytes rest	20.2	Neutrophils rest	0.0
Monocytes LPS	15.7	Colon	0.0
Macrophages rest	19.1	Lung	15.7
Macrophages LPS	0.0	Thymus	8.2
HUVEC none	0.0	Kidney	15.6
HUVEC starved	0.0		

Table BEF. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4145, Run 268861697	Tissue Name	Rel. Exp.(%) Ag4145, Run 268861697
Colon cancer 1	18.8	Bladder cancer NAT 2	13.6
Colon cancer NAT 1	9.0	Bladder cancer NAT 3	0.9
Colon cancer 2	14.5	Bladder cancer NAT 4	8.4
Colon cancer NAT 2	1.7	Adenocarcinoma of the prostate 1	0.0
Colon cancer 3	11.3	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	3.2	Adenocarcinoma of the prostate 3	17.6
Colon malignant cancer 4	17.6	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	4.6	Prostate cancer NAT 5	0.0
Lung cancer 1	22.1	Adenocarcinoma of the prostate 6	1.7
Lung NAT 1	1.5	Adenocarcinoma of the prostate 7	1.0
Lung cancer 2	100.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	3.7	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	11.0	Prostate cancer NAT 10	0.2
Lung NAT 3	4.2	Kidney cancer 1	30.1
metastatic melanoma 1	2.9	KidneyNAT 1	48.0
Melanoma 2	0.2	Kidney cancer 2	47.3

Melanoma 3	3.0	Kidney NAT 2	19.1
metastatic melanoma 4	0.1	Kidney cancer 3	2.9
metastatic melanoma 5	0.0	Kidney NAT 3	1.7
Bladder cancer 1	26.4	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	3.3
Bladder cancer 2	1.1		

CNS_neurodegeneration_v1.0 Summary: Ag4145 This panel confirms the expression of the CG98164-01 gene at low levels in the brain in an independent group of individuals. This gene appears to be slightly upregulated in the temporal cortex of Alzheimer's disease patients. Therefore, therapeutic modulation of the expression or function of this gene may decrease neuronal death and be of use in the treatment of this disease.

General_screening_panel_v1.4 Summary: Ag4145 Highest expression of the CG98164-01 gene is seen in a melanoma cell line (CT=30). Thus, expression of this gene could be used to differentiate between this sample and other samples on this panel and as a marker to detect the presence of melanoma. Furthermore, therapeutic modulation of the expression or function of this gene may be effective in the treatment of melanoma.

In addition, this gene is expressed at much higher levels in fetal lung (CT=32.5) when compared to expression in the adult counterpart (CT=40). Thus, expression of this gene may be used to differentiate between the fetal and adult source of this tissue.

General_screening_panel_v1.5 Summary: Ag4145 Results from one experiment with the CG98164-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 4.1D Summary: Ag4145 Expression of the CG98164-01 gene is restricted to ionomycin and untreated samples from the B cell line Ramos (CTs=34.5). B cells represent a principle component of immunity and contribute to the immune response in a number of important functional roles, including antibody production. Production of antibodies against self-antigens is a major component in autoimmune disorders. Since B cells play an important role in autoimmunity, inflammatory processes and inflammatory cascades, therapeutic modulation of this gene product may reduce or eliminate the symptoms of patients suffering from asthma, allergies, chronic obstructive pulmonary disease, emphysema, Crohn's disease, ulcerative colitis, rheumatoid arthritis, psoriasis, osteoarthritis, systemic lupus erythematosus and other autoimmune disorders. Two

additional experiments with the same probe and primer show low/undetectable levels of expression (CTs>35). (Data not shown.)

General oncology screening panel_v_2.4 Summary: Ag4145 Highest expression of the CG98164-01 gene is detected in lung cancer (CT=29). Significant expression of this gene is seen in number of cancer samples including colon, lung, adenocarcinoma of prostate, bladder and kidney cancers. . Thus, the expression of this gene could be used to distinguish colon, lung and prostate cancers from the normal tissues. Therapeutic modulation of this gene product may be useful in the treatment of these cancers.

BF. CG99588-01: Novel Transmembrane Protein

Expression of gene CG99588-01 was assessed using the primer-probe set Ag4148, described in Table BFA. Results of the RTQ-PCR runs are shown in Tables BFB, BFC, BFD and BFE.

Table BFA. Probe Name Ag4148

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggtcactgtggtgaagagtga-3'	21	12	413
Probe	TET-5'-acccaaactggtgccgttcttcaag-3'-TAMRA	25	36	414
Reverse	5'-cagccagagcacaaaatacac-3'	21	70	415

Table BFB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4148, Run 215309158	Tissue Name	Rel. Exp.(%) Ag4148, Run 215309158
AD 1 Hippo	24.5	Control (Path) 3 Temporal Ctx	9.0
AD 2 Hippo	42.9	Control (Path) 4 Temporal Ctx	49.0
AD 3 Hippo	8.1	AD 1 Occipital Ctx	27.2
AD 4 Hippo	16.7	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	95.3	AD 3 Occipital Ctx	8.6
AD 6 Hippo	72.2	AD 4 Occipital Ctx	29.1
Control 2 Hippo	36.6	AD 5 Occipital Ctx	38.7
Control 4 Hippo	30.8	AD 6 Occipital Ctx	26.1
Control (Path) 3 Hippo	9.9	Control 1 Occipital Ctx	2.9
AD 1 Temporal Ctx	51.8	Control 2 Occipital Ctx	64.6
AD 2 Temporal Ctx	49.0	Control 3 Occipital Ctx	40.6
AD 3 Temporal Ctx	11.4	Control 4 Occipital Ctx	16.8

AD 4 Temporal Ctx	52.5	Control (Path) 1 Occipital Ctx	60.3
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	25.5
AD 5 Sup Temporal Ctx	67.8	Control (Path) 3 Occipital Ctx	3.3
AD 6 Inf Temporal Ctx	65.1	Control (Path) 4 Occipital Ctx	27.0
AD 6 Sup Temporal Ctx	59.9	Control 1 Parietal Ctx	11.2
Control 1 Temporal Ctx	18.9	Control 2 Parietal Ctx	64.6
Control 2 Temporal Ctx	40.6	Control 3 Parietal Ctx	19.5
Control 3 Temporal Ctx	28.3	Control (Path) 1 Parietal Ctx	71.7
Control 3 Temporal Ctx	23.5	Control (Path) 2 Parietal Ctx	34.9
Control (Path) 1 Temporal Ctx	63.3	Control (Path) 3 Parietal Ctx	10.9
Control (Path) 2 Temporal Ctx	53.2	Control (Path) 4 Parietal Ctx	50.0

Table BFC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4148, Run 222181956	Tissue Name	Rel. Exp.(%) Ag4148, Run 222181956
Adipose	19.9	Renal ca. TK-10	7.6
Melanoma* Hs688(A).T	5.9	Bladder	57.4
Melanoma* Hs688(B).T	3.4	Gastric ca. (liver met.) NCI-N87	32.3
Melanoma* M14	6.4	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	4.8	Colon ca. SW-948	3.3
Melanoma* SK-MEL-5	17.2	Colon ca. SW480	8.3
Squamous cell carcinoma SCC-4	6.4	Colon ca.* (SW480 met) SW620	1.7
Testis Pool	12.1	Colon ca. HT29	3.4
Prostate ca.* (bone met) PC-3	13.3	Colon ca. HCT-116	3.6
Prostate Pool	2.6	Colon ca. CaCo-2	1.8
Placenta	37.4	Colon cancer tissue	32.5
Uterus Pool	8.8	Colon ca. SW1116	3.5
Ovarian ca. OVCAR-3	17.8	Colon ca. Colo-205	1.5
Ovarian ca. SK-OV-3	5.9	Colon ca. SW-48	5.8
Ovarian ca. OVCAR-4	8.2	Colon Pool	19.2
Ovarian ca. OVCAR-5	33.7	Small Intestine Pool	22.2
Ovarian ca. IGROV-1	7.8	Stomach Pool	16.5
Ovarian ca. OVCAR-8	22.4	Bone Marrow Pool	10.5
Ovary	58.2	Fetal Heart	6.9
Breast ca. MCF-7	4.5	Heart Pool	11.9

Breast ca. MDA-MB-231	3.6	Lymph Node Pool	14.6
Breast ca. BT 549	1.2	Fetal Skeletal Muscle	3.7
Breast ca. T47D	100.0	Skeletal Muscle Pool	3.8
Breast ca. MDA-N	8.1	Spleen Pool	20.6
Breast Pool	16.4	Thymus Pool	15.6
Trachea	11.5	CNS cancer (glio/astro) U87-MG	7.3
Lung	5.4	CNS cancer (glio/astro) U-118-MG	0.6
Fetal Lung	14.8	CNS cancer (neuro;met) SK-N-AS	26.6
Lung ca. NCI-N417	1.1	CNS cancer (astro) SF-539	8.8
Lung ca. LX-1	6.7	CNS cancer (astro) SNB-75	27.0
Lung ca. NCI-H146	0.2	CNS cancer (glio) SNB-19	4.7
Lung ca. SHP-77	2.8	CNS cancer (glio) SF-295	13.9
Lung ca. A549	1.0	Brain (Amygdala) Pool	15.8
Lung ca. NCI-H526	6.3	Brain (cerebellum)	36.9
Lung ca. NCI-H23	2.5	Brain (fetal)	29.1
Lung ca. NCI-H460	0.5	Brain (Hippocampus) Pool	13.2
Lung ca. HOP-62	1.0	Cerebral Cortex Pool	16.0
Lung ca. NCI-H522	28.7	Brain (Substantia nigra) Pool	19.9
Liver	3.4	Brain (Thalamus) Pool	21.0
Fetal Liver	12.1	Brain (whole)	23.5
Liver ca. HepG2	0.6	Spinal Cord Pool	18.2
Kidney Pool	28.1	Adrenal Gland	32.5
Fetal Kidney	9.3	Pituitary gland Pool	5.2
Renal ca. 786-0	3.1	Salivary Gland	8.4
Renal ca. A498	11.3	Thyroid (female)	10.3
Renal ca. ACHN	10.4	Pancreatic ca. CAPAN2	1.8
Renal ca. UO-31	11.8	Pancreas Pool	24.0

Table BFD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4148, Run 173123939	Tissue Name	Rel. Exp.(%) Ag4148, Run 173123939
Secondary Th1 act	0.3	HUVEC IL-1beta	1.4
Secondary Th2 act	6.4	HUVEC IFN gamma	18.6
Secondary Tr1 act	1.1	HUVEC TNF alpha + IFN gamma	3.1
Secondary Th1 rest	6.7	HUVEC TNF alpha + IL4	1.9
Secondary Th2 rest	3.4	HUVEC IL-11	2.2

Secondary Tr1 rest	11.2	Lung Microvascular EC none	10.8
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	16.6
Primary Th2 act	4.2	Microvascular Dermal EC none	7.3
Primary Tr1 act	1.4	Microvascular Dermal EC TNFalpha + IL-1beta	3.6
Primary Th1 rest	1.1	Bronchial epithelium TNFalpha + IL1beta	6.0
Primary Th2 rest	4.0	Small airway epithelium none	16.7
Primary Tr1 rest	4.2	Small airway epithelium TNFalpha + IL-1beta	12.9
CD45RA CD4 lymphocyte act	1.2	Coronary artery SMC rest	5.7
CD45RO CD4 lymphocyte act	2.7	Coronary artery SMC TNFalpha + IL-1beta	10.5
CD8 lymphocyte act	2.4	Astrocytes rest	9.3
Secondary CD8 lymphocyte rest	2.1	Astrocytes TNFalpha + IL-1beta	1.1
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	0.0
CD4 lymphocyte none	1.4	KU-812 (Basophil) PMA/ionomycin	0.0
2ry Th1/Th2/Tr1_anti-CD95 CH11	8.1	CCD1106 (Keratinocytes) none	8.7
LAK cells rest	17.1	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	2.9
LAK cells IL-2	4.5	Liver cirrhosis	7.3
LAK cells IL-2+IL-12	0.0	NCI-H292 none	20.3
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	15.6
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	28.3
LAK cells PMA/ionomycin	1.0	NCI-H292 IL-13	21.3
NK Cells IL-2 rest	3.7	NCI-H292 IFN gamma	19.5
Two Way MLR 3 day	9.5	HPAEC none	3.4
Two Way MLR 5 day	2.3	HPAEC TNF alpha + IL-1 beta	4.9
Two Way MLR 7 day	0.2	Lung fibroblast none	3.2
PBMC rest	3.6	Lung fibroblast TNF alpha + IL-1 beta	0.0
PBMC PWM	1.7	Lung fibroblast IL-4	1.7
PBMC PHA-L	3.0	Lung fibroblast IL-9	1.1
Ramos (B cell) none	0.0	Lung fibroblast IL-13	1.0

Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	0.0
B lymphocytes PWM	0.0	Dermal fibroblast CCD1070 rest	3.7
B lymphocytes CD40L and IL-4	2.8	Dermal fibroblast CCD1070 TNF alpha	7.2
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	2.8
EOL-1 dbcAMP PMA/ionomycin	2.0	Dermal fibroblast IFN gamma	1.4
Dendritic cells none	71.7	Dermal fibroblast IL-4	0.9
Dendritic cells LPS	8.7	Dermal Fibroblasts rest	1.9
Dendritic cells anti-CD40	81.8	Neutrophils TNFa+LPS	0.0
Monocytes rest	18.8	Neutrophils rest	0.0
Monocytes LPS	8.4	Colon	8.8
Macrophages rest	56.6	Lung	47.3
Macrophages LPS	1.0	Thymus	22.8
HUVEC none	2.2	Kidney	100.0
HUVEC starved	12.2		

Table BFE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4148, Run 268623908	Tissue Name	Rel. Exp.(%) Ag4148, Run 268623908
Colon cancer 1	12.6	Bladder cancer NAT 2	0.3
Colon NAT 1	7.6	Bladder cancer NAT 3	0.1
Colon cancer 2	19.1	Bladder cancer NAT 4	2.1
Colon cancer NAT 2	5.1	Adenocarcinoma of the prostate 1	24.0
Colon cancer 3	23.0	Adenocarcinoma of the prostate 2	1.3
Colon cancer NAT 3	11.9	Adenocarcinoma of the prostate 3	10.6
Colon malignant cancer 4	30.1	Adenocarcinoma of the prostate 4	21.8
Colon normal adjacent tissue 4	4.1	Prostate cancer NAT 5	1.1
Lung cancer 1	13.3	Adenocarcinoma of the prostate 6	5.4
Lung NAT 1	5.0	Adenocarcinoma of the prostate 7	7.9
Lung cancer 2	36.6	Adenocarcinoma of the prostate 8	1.6
Lung NAT 2	3.0	Adenocarcinoma of the prostate	20.0

		9	
Squamous cell carcinoma 3	22.8	Prostate cancer NAT 10	0.4
Lung NAT 3	1.7	Kidney cancer 1	38.4
metastatic melanoma 1	33.9	Kidney NAT 1	24.0
Melanoma 2	8.2	Kidney cancer 2	66.9
Melanoma 3	2.6	Kidney NAT 2	96.6
metastatic melanoma 4	96.6	Kidney cancer 3	17.3
metastatic melanoma 5	100.0	Kidney NAT 3	19.2
Bladder cancer 1	0.8	Kidney cancer 4	8.8
Bladder cancer NAT 1	0.0	Kidney NAT 4	38.4
Bladder cancer 2	4.6		

CNS_neurodegeneration_v1.0 Summary: Ag4148 This panel confirms the expression of the CG99588-01 gene at low levels in the brain in an independent group of individuals. This gene is found to be slightly upregulated in the temporal cortex of Alzheimer's disease patients. Therefore, therapeutic modulation of the expression or function of this gene may decrease neuronal death and be of use in the treatment of this disease.

General_screening_panel_v1.4 Summary: Ag4148 Highest expression of the CG99588-01 gene is detected in breast cancer cell line (CT=30.41). Significant expression of this gene is also seen in cluster of cancer cell lines including CNS, colon, gastric, renal, lung, breast, ovarian, prostate, squamous cell carcinoma, and melanoma cell lines. Therefore, therapeutic modulation of this gene product may be useful in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at low to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at moderate levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Ag4148 Results from one experiment (run 220982871) with the this gene are not included. The amp plot indicates that there were experimental difficulties with this run.

Panel 4.1D Summary: Ag4148 Highest expression of the CG99588-01 gene is detected in kidney (CT=30.3). In addition, moderate to low levels of expression of this gene is also seen primary and secondary Th1, Th2 and Tr1 cells, LAK cells, dendritic cells, monocytes, macrophages, endothelial cells, bronchial and small airway epithelial cells, coronary artery SMC, NCI-H292, astrocytes and normal tissues represent by colon, lung, and thymus. Therefore, therapeutic modulation of the gene product may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4148 Highest expression of the CG99588-01 gene is detected in metastatic melanoma (CT=29). In addition, significant expression of this gene is also seen in number of cancer samples including kidney, colon, adenocarcinoma of prostate, lung and bladder cancer. Therefore, therapeutic modulation of this gene product through the use of small molecule drug may be beneficial in the treatment of these cancers.

20 **BG. CG99618-01: Protein-Tyrosine Phosphatase 2C**

Expression of gene CG99618-01 was assessed using the primer-probe set Ag4151, described in Table BGA. Results of the RTQ-PCR runs are shown in Tables BGB, BGC and BGD.

Table BGA. Probe Name Ag4151

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-catcatgggcaattaaaagaga-3'	22	258	416
Probe	TET-5'-aaaatcctctgaactgtgcagatcct-3'-TAMRA	26	304	417
Reverse	5'-catgaaaccacctttgagaagt-3'	22	330	418

25 **Table BGB. CNS_neurodegeneration_v1.0**

Tissue Name	Rel. Exp.(%) Ag4151, Run 215318747	Tissue Name	Rel. Exp.(%) Ag4151, Run 215318747
AD 1 Hippo	3.7	Control (Path) 3 Temporal Ctx	2.0

AD 2 Hippo	6.1	Control (Path) 4 Temporal Ctx	30.8
AD 3 Hippo	0.7	AD 1 Occipital Ctx	12.9
AD 4 Hippo	3.5	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	100.0	AD 3 Occipital Ctx	3.8
AD 6 Hippo	12.4	AD 4 Occipital Ctx	15.9
Control 2 Hippo	7.3	AD 5 Occipital Ctx	8.2
Control 4 Hippo	4.0	AD 6 Occipital Ctx	17.2
Control (Path) 3 Hippo	0.0	Control 1 Occipital Ctx	1.1
AD 1 Temporal Ctx	3.7	Control 2 Occipital Ctx	32.1
AD 2 Temporal Ctx	14.1	Control 3 Occipital Ctx	7.3
AD 3 Temporal Ctx	3.2	Control 4 Occipital Ctx	4.9
AD 4 Temporal Ctx	17.6	Control (Path) 1 Occipital Ctx	49.3
AD 5 Inf Temporal Ctx	90.8	Control (Path) 2 Occipital Ctx	11.3
AD 5 Sup Temporal Ctx	16.8	Control (Path) 3 Occipital Ctx	0.0
AD 6 Inf Temporal Ctx	0.0	Control (Path) 4 Occipital Ctx	19.1
AD 6 Sup Temporal Ctx	37.6	Control 1 Parietal Ctx	2.8
Control 1 Temporal Ctx	4.8	Control 2 Parietal Ctx	26.2
Control 2 Temporal Ctx	8.5	Control 3 Parietal Ctx	12.4
Control 3 Temporal Ctx	7.9	Control (Path) 1 Parietal Ctx	35.8
Control 4 Temporal Ctx	11.2	Control (Path) 2 Parietal Ctx	13.0
Control (Path) 1 Temporal Ctx	28.1	Control (Path) 3 Parietal Ctx	0.0
Control (Path) 2 Temporal Ctx	31.4	Control (Path) 4 Parietal Ctx	31.0

Table BGC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4151, Run 221034281	Tissue Name	Rel. Exp.(%) Ag4151, Run 221034281
Adipose	0.5	Renal ca. TK-10	0.0
Melanoma* Hs688(A).T	0.9	Bladder	1.5
Melanoma* Hs688(B).T	0.0	Gastric ca. (liver met.) NCI-N87	0.8
Melanoma* M14	0.0	Gastric ca. KATO III	0.0
Melanoma* LOXIMVI	0.0	Colon ca. SW-948	0.2
Melanoma* SK-MEL-5	100.0	Colon ca. SW480	0.0
Squamous cell carcinoma SCC-4	0.0	Colon ca.* (SW480 met) SW620	0.0
Testis Pool	4.2	Colon ca. HT29	0.0
Prostate ca.* (bone met) PC-3	0.4	Colon ca. HCT-116	0.3

Prostate Pool	4.2	Colon ca. CaCo-2	0.1
Placenta	1.5	Colon cancer tissue	0.4
Uterus Pool	0.7	Colon ca. SW1116	0.1
Ovarian ca. OVCAR-3	1.1	Colon ca. Colo-205	0.0
Ovarian ca. SK-OV-3	0.5	Colon ca. SW-48	0.0
Ovarian ca. OVCAR-4	0.5	Colon Pool	4.2
Ovarian ca. OVCAR-5	0.4	Small Intestine Pool	2.2
Ovarian ca. IGROV-1	0.0	Stomach Pool	0.9
Ovarian ca. OVCAR-8	1.5	Bone Marrow Pool	1.0
Ovary	3.0	Fetal Heart	4.5
Breast ca. MCF-7	0.1	Heart Pool	1.0
Breast ca. MDA-MB-231	0.5	Lymph Node Pool	2.1
Breast ca. BT 549	3.5	Fetal Skeletal Muscle	2.3
Breast ca. T47D	0.2	Skeletal Muscle Pool	1.6
Breast ca. MDA-N	0.2	Spleen Pool	1.1
Breast Pool	1.6	Thymus Pool	4.0
Trachea	1.9	CNS cancer (glio/astro) U87-MG	0.8
Lung	5.6	CNS cancer (glio/astro) U-118-MG	1.4
Fetal Lung	4.1	CNS cancer (neuro;met) SK-N-AS	0.2
Lung ca. NCI-N417	0.0	CNS cancer (astro) SF-539	0.0
Lung ca. LX-1	0.0	CNS cancer (astro) SNB-75	5.4
Lung ca. NCI-H146	0.0	CNS cancer (glio) SNB-19	0.5
Lung ca. SHP-77	0.2	CNS cancer (glio) SF-295	2.3
Lung ca. A549	0.2	Brain (Amygdala) Pool	4.3
Lung ca. NCI-H526	0.0	Brain (cerebellum)	0.9
Lung ca. NCI-H23	1.2	Brain (fetal)	2.6
Lung ca. NCI-H460	0.5	Brain (Hippocampus) Pool	3.8
Lung ca. HOP-62	1.5	Cerebral Cortex Pool	8.5
Lung ca. NCI-H522	0.0	Brain (Substantia nigra) Pool	3.9
Liver	0.0	Brain (Thalamus) Pool	13.6
Fetal Liver	0.2	Brain (whole)	2.4
Liver ca. HepG2	0.0	Spinal Cord Pool	2.5
Kidney Pool	4.7	Adrenal Gland	1.5
Fetal Kidney	2.3	Pituitary gland Pool	1.9
Renal ca. 786-0	2.2	Salivary Gland	0.2
Renal ca. A498	0.0	Thyroid (female)	0.0
Renal ca. ACHN	1.1	Pancreatic ca. CAPAN2	0.5
Renal ca. UO-31	1.0	Pancreas Pool	2.1

Table BGD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4151, Run 173124788	Tissue Name	Rel. Exp.(%) Ag4151, Run 173124788
Secondary Th1 act	0.0	HUVEC IL-1beta	0.0
Secondary Th2 act	4.6	HUVEC IFN gamma	2.6
Secondary Tr1 act	0.9	HUVEC TNF alpha + IFN gamma	0.0
Secondary Th1 rest	0.0	HUVEC TNF alpha + IL4	0.0
Secondary Th2 rest	0.0	HUVEC IL-11	0.0
Secondary Tr1 rest	3.2	Lung Microvascular EC none	3.3
Primary Th1 act	0.0	Lung Microvascular EC TNFalpha + IL-1beta	0.0
Primary Th2 act	0.0	Microvascular Dermal EC none	0.0
Primary Tr1 act	0.0	Microvascular Dermal EC TNFalpha + IL-1beta	0.0
Primary Th1 rest	0.0	Bronchial epithelium TNFalpha + IL1beta	16.5
Primary Th2 rest	0.0	Small airway epithelium none	1.0
Primary Tr1 rest	0.0	Small airway epithelium TNFalpha + IL-1beta	0.0
CD45RA CD4 lymphocyte act	0.0	Coronary artery SMC rest	0.0
CD45RO CD4 lymphocyte act	0.0	Coronary artery SMC TNFalpha + IL-1beta	0.0
CD8 lymphocyte act	0.0	Astrocytes rest	6.1
Secondary CD8 lymphocyte rest	0.0	Astrocytes TNFalpha + IL-1beta	2.9
Secondary CD8 lymphocyte act	0.0	KU-812 (Basophil) rest	2.0
CD4 lymphocyte none	0.0	KU-812 (Basophil) PMA/ionomycin	3.4
2ry Th1/Th2/Tr1_anti-CD95 CH11	2.3	CCD1106 (Keratinocytes) none	1.8
LAK cells rest	0.0	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	32.1
LAK cells IL-2	4.0	Liver cirrhosis	0.0
LAK cells IL-2+IL-12	0.0	NCI-H292 none	1.8
LAK cells IL-2+IFN gamma	0.0	NCI-H292 IL-4	3.2
LAK cells IL-2+ IL-18	0.0	NCI-H292 IL-9	1.8
LAK cells PMA/ionomycin	0.0	NCI-H292 IL-13	0.0

NK Cells IL-2 rest	0.0	NCI-H292 IFN gamma	9.3
Two Way MLR 3 day	0.5	HPAEC none	0.0
Two Way MLR 5 day	0.0	HPAEC TNF alpha + IL-1 beta	0.0
Two Way MLR 7 day	0.0	Lung fibroblast none	5.8
PBMC rest	2.0	Lung fibroblast TNF alpha + IL-1 beta	4.1
PBMC PWM	0.0	Lung fibroblast IL-4	0.0
PBMC PHA-L	0.0	Lung fibroblast IL-9	0.0
Ramos (B cell) none	0.0	Lung fibroblast IL-13	4.6
Ramos (B cell) ionomycin	0.0	Lung fibroblast IFN gamma	1.7
B lymphocytes PWM	5.5	Dermal fibroblast CCD1070 rest	12.4
B lymphocytes CD40L and IL-4	0.0	Dermal fibroblast CCD1070 TNF alpha	2.3
EOL-1 dbcAMP	0.0	Dermal fibroblast CCD1070 IL-1 beta	0.0
EOL-1 dbcAMP PMA/ionomycin	11.1	Dermal fibroblast IFN gamma	0.0
Dendritic cells none	0.0	Dermal fibroblast IL-4	2.9
Dendritic cells LPS	0.0	Dermal Fibroblasts rest	0.0
Dendritic cells anti-CD40	0.0	Neutrophils TNFa+LPS	0.0
Monocytes rest	2.6	Neutrophils rest	0.0
Monocytes LPS	0.6	Colon	0.0
Macrophages rest	0.0	Lung	3.1
Macrophages LPS	0.0	Thymus	14.6
HUVEC none	0.0	Kidney	100.0
HUVEC starved	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4151 This panel confirms the expression of the CG99618-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4151 Highest expression of the CG99618-01 gene is detected in melanoma SK-MEL-5 cell line (CT=29.6). Therefore, expression of this gene can be used to distinguish this sample from other samples in the panel. Low expression of this gene is also detected in a breast cancer and a CNS cancer

cell lines. Therefore, therapeutic modulation of this gene product may be useful in the treatment of melanoma, breast and CNS cancers.

In addition, this gene is expressed at low levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4151 Highest expression of the CG99618-01 gene is detected in kidney (CT=32.4). Therefore, expression of this gene can be used to distinguish kidney from other samples in this panel. In addition, therapeutic modulation of this gene product may be useful in the treatment of inflammatory and autoimmune diseases that affect kidney such as lupus and glomerulonephritis.

Low levels of expression of this gene is also seen in TNFalpha + IL-1beta treated keratinocytes. Interestingly, this expression in treated cells is higher (CT=34) as compared to the untreated keratinocytes (CT=38). Therefore, expression of this gene can be used to distinguish the treated from untreated keratinocytes. In addition, therapeutic modulation of this gene product may be useful in the treatment of psoriasis and wound healing.

BH. CG99832-01: Novel Gene Containing NUDIX Hydrolase Domain

Expression of gene CG99832-01 was assessed using the primer-probe set Ag4157, described in Table BHA. Results of the RTQ-PCR runs are shown in Tables BHB, BHC, BHD and BHE.

Table BHA. Probe Name Ag4157

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agcttgctcgtttgatcatcat-3'	22	543	419
Probe	TET-5'-tccaggaattccaaaagacacaaaat-3'-TAMRA	26	565	420
Reverse	5'-cactcaatgttccgaatttctc-3'	22	609	421

Table BHB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4157, Run 215331739	Tissue Name	Rel. Exp.(%) Ag4157, Run 215331739
AD 1 Hippo	11.7	Control (Path) 3 Temporal Ctx	5.7
AD 2 Hippo	24.8	Control (Path) 4 Temporal Ctx	37.1

AD 3 Hippo	7.5	AD 1 Occipital Ctx	15.1
AD 4 Hippo	5.7	AD 2 Occipital Ctx (Missing)	0.0
AD 5 hippo	78.5	AD 3 Occipital Ctx	4.0
AD 6 Hippo	52.9	AD 4 Occipital Ctx	15.6
Control 2 Hippo	28.3	AD 5 Occipital Ctx	31.6
Control 4 Hippo	8.2	AD 6 Occipital Ctx	41.2
Control (Path) 3 Hippo	10.6	Control 1 Occipital Ctx	3.8
AD 1 Temporal Ctx	14.7	Control 2 Occipital Ctx	52.1
AD 2 Temporal Ctx	31.2	Control 3 Occipital Ctx	12.4
AD 3 Temporal Ctx	4.7	Control 4 Occipital Ctx	6.7
AD 4 Temporal Ctx	20.4	Control (Path) 1 Occipital Ctx	79.6
AD 5 Inf Temporal Ctx	100.0	Control (Path) 2 Occipital Ctx	9.7
AD 5 Sup Temporal Ctx	40.6	Control (Path) 3 Occipital Ctx	4.5
AD 6 Inf Temporal Ctx	63.3	Control (Path) 4 Occipital Ctx	13.2
AD 6 Sup Temporal Ctx	66.0	Control 1 Parietal Ctx	6.4
Control 1 Temporal Ctx	5.6	Control 2 Parietal Ctx	38.7
Control 2 Temporal Ctx	40.6	Control 3 Parietal Ctx	17.2
Control 3 Temporal Ctx	15.5	Control (Path) 1 Parietal Ctx	79.6
Control 4 Temporal Ctx	5.0	Control (Path) 2 Parietal Ctx	28.9
Control (Path) 1 Temporal Ctx	70.7	Control (Path) 3 Parietal Ctx	3.6
Control (Path) 2 Temporal Ctx	34.6	Control (Path) 4 Parietal Ctx	40.9

Table BHC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4157, Run 221117670	Tissue Name	Rel. Exp.(%) Ag4157, Run 221117670
Adipose	18.2	Renal ca. TK-10	28.1
Melanoma* Hs688(A).T	14.1	Bladder	23.8
Melanoma* Hs688(B).T	12.1	Gastric ca. (liver met.) NCI-N87	23.5
Melanoma* M14	32.3	Gastric ca. KATO III	24.0
Melanoma* LOXIMVI	21.5	Colon ca. SW-948	6.1
Melanoma* SK-MEL-5	41.8	Colon ca. SW480	21.8
Squamous cell carcinoma SCC-4	11.7	Colon ca.* (SW480 met) SW620	14.4
Testis Pool	8.3	Colon ca. HT29	18.4
Prostate ca.* (bone met) PC-3	13.5	Colon ca. HCT-116	51.8
Prostate Pool	9.4	Colon ca. CaCo-2	80.1

Placenta	32.8	Colon cancer tissue	14.7
Uterus Pool	8.3	Colon ca. SW1116	2.7
Ovarian ca. OVCAR-3	31.2	Colon ca. Colo-205	3.7
Ovarian ca. SK-OV-3	25.3	Colon ca. SW-48	4.5
Ovarian ca. OVCAR-4	14.3	Colon Pool	17.3
Ovarian ca. OVCAR-5	32.5	Small Intestine Pool	17.3
Ovarian ca. IGROV-1	15.8	Stomach Pool	8.9
Ovarian ca. OVCAR-8	11.6	Bone Marrow Pool	10.2
Ovary	11.4	Fetal Heart	12.5
Breast ca. MCF-7	53.6	Heart Pool	7.6
Breast ca. MDA-MB-231	34.6	Lymph Node Pool	17.7
Breast ca. BT 549	34.6	Fetal Skeletal Muscle	9.8
Breast ca. T47D	100.0	Skeletal Muscle Pool	8.8
Breast ca. MDA-N	16.7	Spleen Pool	16.4
Breast Pool	18.6	Thymus Pool	18.4
Trachea	11.0	CNS cancer (glio/astro) U87-MG	15.3
Lung	7.4	CNS cancer (glio/astro) U-118-MG	20.6
Fetal Lung	45.7	CNS cancer (neuro;met) SK-N-AS	27.2
Lung ca. NCI-N417	9.6	CNS cancer (astro) SF-539	15.6
Lung ca. LX-1	24.1	CNS cancer (astro) SNB-75	31.4
Lung ca. NCI-H146	17.1	CNS cancer (glio) SNB-19	17.2
Lung ca. SHP-77	48.3	CNS cancer (glio) SF-295	21.5
Lung ca. A549	14.1	Brain (Amygdala) Pool	10.5
Lung ca. NCI-H526	13.7	Brain (cerebellum)	18.3
Lung ca. NCI-H23	32.1	Brain (fetal)	48.3
Lung ca. NCI-H460	10.8	Brain (Hippocampus) Pool	13.7
Lung ca. HOP-62	18.3	Cerebral Cortex Pool	17.9
Lung ca. NCI-H522	36.3	Brain (Substantia nigra) Pool	12.6
Liver	1.7	Brain (Thalamus) Pool	20.7
Fetal Liver	43.2	Brain (whole)	19.2
Liver ca. HepG2	17.8	Spinal Cord Pool	6.2
Kidney Pool	27.0	Adrenal Gland	16.2
Fetal Kidney	55.5	Pituitary gland Pool	4.5
Renal ca. 786-0	27.9	Salivary Gland	5.4
Renal ca. A498	4.9	Thyroid (female)	2.8
Renal ca. ACHN	8.5	Pancreatic ca. CAPAN2	30.4

Renal ca. UO-31	12.2	Pancreas Pool	16.8
-----------------	------	---------------	------

Table BHD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4157, Run 173123943	Tissue Name	Rel. Exp.(%) Ag4157, Run 173123943
Secondary Th1 act	55.1	HUVEC IL-1 beta	20.4
Secondary Th2 act	54.7	HUVEC IFN gamma	28.7
Secondary Tr1 act	42.9	HUVEC TNF alpha + IFN gamma	15.6
Secondary Th1 rest	23.7	HUVEC TNF alpha + IL4	12.9
Secondary Th2 rest	33.9	HUVEC IL-11	24.8
Secondary Tr1 rest	23.7	Lung Microvascular EC none	24.7
Primary Th1 act	29.7	Lung Microvascular EC TNFalpha + IL-1 beta	19.5
Primary Th2 act	44.8	Microvascular Dermal EC none	11.1
Primary Tr1 act	36.1	Microvascular Dermal EC TNFalpha + IL-1 beta	11.7
Primary Th1 rest	38.4	Bronchial epithelium TNFalpha + IL1 beta	19.9
Primary Th2 rest	44.4	Small airway epithelium none	4.5
Primary Tr1 rest	48.0	Small airway epithelium TNFalpha + IL-1 beta	12.9
CD45RA CD4 lymphocyte act	21.2	Coronary artery SMC rest	9.3
CD45RO CD4 lymphocyte act	46.3	Coronary artery SMC TNFalpha + IL-1 beta	11.2
CD8 lymphocyte act	28.7	Astrocytes rest	11.1
Secondary CD8 lymphocyte rest	27.7	Astrocytes TNFalpha + IL-1 beta	10.9
Secondary CD8 lymphocyte act	27.7	KU-812 (Basophil) rest	39.0
CD4 lymphocyte none	16.8	KU-812 (Basophil) PMA/ionomycin	52.9
2ry Th1/Th2/Tr1_anti-CD95 CH11	52.1	CCD1106 (Keratinocytes) none	13.0
LAK cells rest	34.2	CCD1106 (Keratinocytes) TNFalpha + IL-1 beta	14.3
LAK cells IL-2	37.4	Liver cirrhosis	7.2
LAK cells IL-2+IL-12	22.4	NCI-H292 none	11.2
LAK cells IL-2+IFN gamma	36.1	NCI-H292 IL-4	23.8
LAK cells IL-2+ IL-18	41.8	NCI-H292 IL-9	21.3

LAK cells PMA/ionomycin	39.8	NCI-H292 IL-13	21.9
NK Cells IL-2 rest	54.7	NCI-H292 IFN gamma	16.7
Two Way MLR 3 day	43.2	HPAEC none	19.3
Two Way MLR 5 day	44.8	HPAEC TNF alpha + IL-1 beta	35.4
Two Way MLR 7 day	35.4	Lung fibroblast none	7.6
PBMC rest	20.2	Lung fibroblast TNF alpha + IL-1 beta	7.7
PBMC PWM	22.5	Lung fibroblast IL-4	7.3
PBMC PHA-L	33.7	Lung fibroblast IL-9	9.5
Ramos (B cell) none	75.8	Lung fibroblast IL-13	8.5
Ramos (B cell) ionomycin	64.2	Lung fibroblast IFN gamma	7.6
B lymphocytes PWM	35.4	Dermal fibroblast CCD1070 rest	17.6
B lymphocytes CD40L and IL-4	48.6	Dermal fibroblast CCD1070 TNF alpha	52.5
EOL-1 dbcAMP	54.3	Dermal fibroblast CCD1070 IL-1 beta	13.5
EOL-1 dbcAMP PMA/ionomycin	86.5	Dermal fibroblast IFN gamma	15.8
Dendritic cells none	58.6	Dermal fibroblast IL-4	18.4
Dendritic cells LPS	40.6	Dermal Fibroblasts rest	6.4
Dendritic cells anti- CD40	50.3	Neutrophils TNFa+LPS	28.7
Monocytes rest	52.1	Neutrophils rest	87.1
Monocytes LPS	100.0	Colon	6.9
Macrophages rest	38.4	Lung	11.6
Macrophages LPS	25.2	Thymus	66.4
HUVEC none	15.8	Kidney	19.9
HUVEC starved	21.0		

Table BHE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4157, Run 268624005	Tissue Name	Rel. Exp.(%) Ag4157, Run 268624005
Colon cancer 1	12.8	Bladder cancer NAT 2	0.4
Colon NAT 1	4.8	Bladder cancer NAT 3	1.8
Colon cancer 2	26.1	Bladder cancer NAT 4	6.6
Colon cancer NAT 2	5.6	Adenocarcinoma of the prostate 1	34.9
Colon cancer 3	24.8	Adenocarcinoma of the prostate 2	3.5
Colon cancer NAT 3	13.2	Adenocarcinoma of the prostate 3	9.0

Colon malignant cancer 4	34.4	Adenocarcinoma of the prostate 4	34.4
Colon normal adjacent tissue 4	5.0	Prostate cancer NAT 5	3.6
Lung cancer 1	11.7	Adenocarcinoma of the prostate 6	3.8
Lung NAT 1	2.5	Adenocarcinoma of the prostate 7	5.4
Lung cancer 2	66.9	Adenocarcinoma of the prostate 8	1.9
Lung NAT 2	4.0	Adenocarcinoma of the prostate 9	17.3
Squamous cell carcinoma 3	18.7	Prostate cancer NAT 10	3.3
Lung NAT 3	1.1	Kidney cancer 1	15.2
metastatic melanoma 1	23.0	Kidney NAT 1	8.5
Melanoma 2	2.1	Kidney cancer 2	67.4
Melanoma 3	5.3	Kidney NAT 2	11.2
metastatic melanoma 4	92.0	Kidney cancer 3	12.5
metastatic melanoma 5	100.0	Kidney NAT 3	5.0
Bladder cancer 1	2.1	Kidney cancer 4	16.8
Bladder cancer NAT 1	0.0	Kidney NAT 4	4.2
Bladder cancer 2	7.1		

CNS_neurodegeneration_v1.0 Summary: Ag4157 This panel confirms the expression of the CG99832-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4157 Highest expression of the CG99832-01 gene is detected in breast cancer T47D cell line (CT=29). Moderate levels of expression of this gene is seen in cluster of cancer lines including pancreatic, CNS, colon, gastric, renal, lung, breast, ovarian, prostate, squamous cell carcinoma, and melanoma cell lines. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, this gene is expressed at much higher levels in fetal (CT=30) when compared to adult liver (CT=34.8). This observation suggests that expression of this gene can be used to distinguish fetal from adult liver. In addition, the relative overexpression of this gene in fetal liver suggests that the protein product may enhance liver or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of liver related diseases.

In addition, this gene is expressed at moderate levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4157 Highest expression of the CG99832-01 gene is detected in LPS treated monocytes (CT=29). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

General oncology screening panel_v_2.4 Summary: Ag4157 Highest expression of the CG99832-01 gene is detected in metastatic melanoma (CT=29). Higher expression of this gene is seen in cancer samples including colon cancer, kidney cancer, lung cancer, prostate adenocarcinoma and melanoma. Therefore, therapeutic modulation of this gene may be useful in the treatment of these cancers.

BI. CG99842-01: Tensin-Like

Expression of gene CG99842-01 was assessed using the primer-probe set Ag4158, described in Table BIA. Results of the RTQ-PCR runs are shown in Tables BIB, BIC and
5 BID.

Table BIA. Probe Name Ag4158

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-aagggataccaagtgagaaagc-3'	22	221	422
Probe	TET-5'-ccttcagttaacaaaggggtacatca-3'-TAMRA	27	245	423
Reverse	5'-tggtccaattgtcacctgattt-3'	22	293	424

Table BIB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4158, Run 215337521	Tissue Name	Rel. Exp.(%) Ag4158, Run 215337521
AD 1 Hippo	10.7	Control (Path) 3 Temporal Ctx	11.7
AD 2 Hippo	28.3	Control (Path) 4 Temporal Ctx	36.3
AD 3 Hippo	13.4	AD 1 Occipital Ctx	23.0
AD 4 Hippo	15.3	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	84.7	AD 3 Occipital Ctx	9.2
AD 6 Hippo	92.7	AD 4 Occipital Ctx	13.4
Control 2 Hippo	26.2	AD 5 Occipital Ctx	20.4
Control 4 Hippo	25.5	AD 6 Occipital Ctx	27.9
Control (Path) 3 Hippo	11.8	Control 1 Occipital Ctx	11.2
AD 1 Temporal Ctx	14.8	Control 2 Occipital Ctx	39.2
AD 2 Temporal Ctx	35.1	Control 3 Occipital Ctx	31.2
AD 3 Temporal Ctx	11.1	Control 4 Occipital Ctx	7.6
AD 4 Temporal Ctx	16.0	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	87.1	Control (Path) 2 Occipital Ctx	15.2
AD 5 Sup Temporal Ctx	44.1	Control (Path) 3 Occipital Ctx	6.9
AD 6 Inf Temporal Ctx	91.4	Control (Path) 4 Occipital Ctx	30.8
AD 6 Sup Temporal Ctx	89.5	Control 1 Parietal Ctx	8.2
Control 1 Temporal Ctx	10.8	Control 2 Parietal Ctx	48.3
Control 2 Temporal Ctx	13.8	Control 3 Parietal Ctx	4.9
Control 3 Temporal Ctx	17.7	Control (Path) 1 Parietal Ctx	37.1
Control 3 Temporal Ctx	7.6	Control (Path) 2 Parietal Ctx	30.8
Control (Path) 1	44.8	Control (Path) 3 Parietal Ctx	13.1

Temporal Ctx			
Control (Path) 2 Temporal Ctx	45.1	Control (Path) 4 Parietal Ctx	54.0

Table BIC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4158, Run 221117871	Tissue Name	Rel. Exp.(%) Ag4158, Run 221117871
Adipose	8.5	Renal ca. TK-10	32.1
Melanoma* Hs688(A).T	16.7	Bladder	26.4
Melanoma* Hs688(B).T	16.3	Gastric ca. (liver met.) NCI-N87	56.6
Melanoma* M14	11.9	Gastric ca. KATO III	54.7
Melanoma* LOXIMVI	30.1	Colon ca. SW-948	11.4
Melanoma* SK-MEL-5	29.9	Colon ca. SW480	36.3
Squamous cell carcinoma SCC-4	12.8	Colon ca.* (SW480 met) SW620	38.7
Testis Pool	5.0	Colon ca. HT29	18.7
Prostate ca.* (bone met) PC-3	31.4	Colon ca. HCT-116	100.0
Prostate Pool	7.5	Colon ca. CaCo-2	31.0
Placenta	3.5	Colon cancer tissue	18.3
Uterus Pool	6.7	Colon ca. SW1116	4.2
Ovarian ca. OVCAR-3	26.8	Colon ca. Colo-205	4.1
Ovarian ca. SK-OV-3	36.6	Colon ca. SW-48	4.0
Ovarian ca. OVCAR-4	6.4	Colon Pool	15.7
Ovarian ca. OVCAR-5	36.6	Small Intestine Pool	18.4
Ovarian ca. IGROV-1	9.5	Stomach Pool	11.0
Ovarian ca. OVCAR-8	5.0	Bone Marrow Pool	9.2
Ovary	9.7	Fetal Heart	5.2
Breast ca. MCF-7	48.6	Heart Pool	6.7
Breast ca. MDA-MB- 231	34.6	Lymph Node Pool	17.7
Breast ca. BT 549	36.6	Fetal Skeletal Muscle	5.1
Breast ca. T47D	52.9	Skeletal Muscle Pool	11.2
Breast ca. MDA-N	10.4	Spleen Pool	15.4
Breast Pool	13.5	Thymus Pool	14.3
Trachea	13.8	CNS cancer (glio/astro) U87-MG	34.4
Lung	10.5	CNS cancer (glio/astro) U-118- MG	62.4
Fetal Lung	24.5	CNS cancer (neuro;met) SK-N-AS	28.7

Lung ca. NCI-N417	1.9	CNS cancer (astro) SF-539	10.2
Lung ca. LX-1	50.0	CNS cancer (astro) SNB-75	28.1
Lung ca. NCI-H146	11.1	CNS cancer (glio) SNB-19	6.6
Lung ca. SHP-77	38.4	CNS cancer (glio) SF-295	41.2
Lung ca. A549	39.0	Brain (Amygdala) Pool	7.8
Lung ca. NCI-H526	1.8	Brain (cerebellum)	8.2
Lung ca. NCI-H23	40.3	Brain (fetal)	18.6
Lung ca. NCI-H460	40.9	Brain (Hippocampus) Pool	8.7
Lung ca. HOP-62	10.5	Cerebral Cortex Pool	7.2
Lung ca. NCI-H522	17.6	Brain (Substantia nigra) Pool	4.8
Liver	0.0	Brain (Thalamus) Pool	14.6
Fetal Liver	12.5	Brain (whole)	6.4
Liver ca. HepG2	18.3	Spinal Cord Pool	9.5
Kidney Pool	33.7	Adrenal Gland	12.5
Fetal Kidney	26.1	Pituitary gland Pool	3.9
Renal ca. 786-0	20.4	Salivary Gland	4.1
Renal ca. A498	3.5	Thyroid (female)	2.2
Renal ca. ACHN	28.5	Pancreatic ca. CAPAN2	50.0
Renal ca. UO-31	19.8	Pancreas Pool	22.4

Table BID. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4158, Run 173123945	Tissue Name	Rel. Exp.(%) Ag4158, Run 173123945
Secondary Th1 act	46.0	HUVEC IL-1beta	17.3
Secondary Th2 act	63.3	HUVEC IFN gamma	14.9
Secondary Tr1 act	35.4	HUVEC TNF alpha + IFN gamma	14.3
Secondary Th1 rest	7.0	HUVEC TNF alpha + IL4	15.8
Secondary Th2 rest	13.2	HUVEC IL-11	12.2
Secondary Tr1 rest	11.7	Lung Microvascular EC none	22.5
Primary Th1 act	55.9	Lung Microvascular EC TNFalpha + IL-1beta	27.4
Primary Th2 act	81.8	Microvascular Dermal EC none	15.7
Primary Tr1 act	51.1	Microvascular Dermal EC TNFalpha + IL-1beta	10.0
Primary Th1 rest	9.5	Bronchial epithelium TNFalpha + IL1beta	12.5
Primary Th2 rest	5.7	Small airway epithelium none	5.2
Primary Tr1 rest	20.2	Small airway epithelium TNFalpha + IL-1beta	18.0

CD45RA CD4 lymphocyte act	31.2	Coronary artery SMC rest	3.3
CD45RO CD4 lymphocyte act	52.5	Coronary artery SMC TNFalpha + IL-1beta	6.9
CD8 lymphocyte act	43.8	Astrocytes rest	9.7
Secondary CD8 lymphocyte rest	40.3	Astrocytes TNFalpha + IL-1beta	7.1
Secondary CD8 lymphocyte act	13.7	KU-812 (Basophil) rest	16.8
CD4 lymphocyte none	13.0	KU-812 (Basophil) PMA/ionomycin	14.8
2ry Th1/Th2/Tr1_anti-CD95 CH11	20.6	CCD1106 (Keratinocytes) none	21.8
LAK cells rest	12.6	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	12.0
LAK cells IL-2	21.8	Liver cirrhosis	1.4
LAK cells IL-2+IL-12	20.3	NCI-H292 none	21.5
LAK cells IL-2+IFN gamma	26.2	NCI-H292 IL-4	33.4
LAK cells IL-2+ IL-18	26.6	NCI-H292 IL-9	43.2
LAK cells PMA/ionomycin	42.0	NCI-H292 IL-13	31.2
NK Cells IL-2 rest	33.7	NCI-H292 IFN gamma	30.8
Two Way MLR 3 day	32.1	HPAEC none	11.1
Two Way MLR 5 day	24.8	HPAEC TNF alpha + IL-1 beta	16.8
Two Way MLR 7 day	18.8	Lung fibroblast none	12.3
PBMC rest	8.1	Lung fibroblast TNF alpha + IL-1 beta	4.6
PBMC PWM	30.4	Lung fibroblast IL-4	8.2
PBMC PHA-L	27.5	Lung fibroblast IL-9	13.4
Ramos (B cell) none	59.5	Lung fibroblast IL-13	11.8
Ramos (B cell) ionomycin	45.4	Lung fibroblast IFN gamma	15.5
B lymphocytes PWM	25.3	Dermal fibroblast CCD1070 rest	25.0
B lymphocytes CD40L and IL-4	24.0	Dermal fibroblast CCD1070 TNF alpha	34.9
EOL-1 dbcAMP	29.5	Dermal fibroblast CCD1070 IL-1 beta	9.7
EOL-1 dbcAMP PMA/ionomycin	18.6	Dermal fibroblast IFN gamma	9.3
Dendritic cells none	12.1	Dermal fibroblast IL-4	17.8
Dendritic cells LPS	9.9	Dermal Fibroblasts rest	3.9

Dendritic cells anti-CD40	14.0	Neutrophils TNFa+LPS	6.7
Monocytes rest	15.0	Neutrophils rest	8.4
Monocytes LPS	27.5	Colon	6.4
Macrophages rest	16.7	Lung	6.7
Macrophages LPS	4.4	Thymus	34.6
HUVEC none	9.2	Kidney	100.0
HUVEC starved	13.9		

CNS_neurodegeneration_v1.0 Summary: Ag4158 This panel confirms the expression of the CG99842-01 gene at low levels in the brains of an independent group of individuals. However, no differential expression of this gene was detected between Alzheimer's diseased postmortem brains and those of non-demented controls in this experiment. Please see Panel 1.4 for a discussion of the potential utility of this gene in treatment of central nervous system disorders.

General_screening_panel_v1.4 Summary: Ag4158 Highest expression of the CG99842-01 gene is detected in colon cancer HCT-116 cell line (CT=30.7). Moderate expression of this gene is associated with cluster of cancer cell lines including pancreatic, CNS, colon, gastric, renal, lung, breast, ovarian, prostate, squamous cell carcinoma and melanoma cancer cell lines. Therefore, therapeutic modulation of this gene product may be useful in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at low to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, fetal liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

Interestingly, this gene is expressed at much higher levels in fetal (CT=33.7) when compared to adult liver (CT=40). This observation suggests that expression of this gene can be used to distinguish fetal from adult liver. In addition, the relative overexpression of this gene in fetal liver suggests that the protein product may enhance liver growth or development in the fetus and thus may also act in a regenerative capacity in the adult. Therefore, therapeutic modulation of the protein encoded by this gene could be useful in treatment of liver related diseases.

In addition, this gene is expressed at low to moderate levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra,

thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4158 Highest expression of the CG99842-01 gene is detected in kidney (CT=31). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon, lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in General_screening_panel_v1.4 and also suggests a role for the gene product in cell survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

20 BJ. CG99944-01: ABC Transporter

Expression of gene CG99944-01 was assessed using the primer-probe set Ag4184, described in Table BJA. Results of the RTQ-PCR runs are shown in Table BJB.

Table BJA. Probe Name Ag4184

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-agctaaaaggggaagacatcac-3'	22	965	425
Probe	TET-5'-aaaacctcgaaagatcactgaacatg-3'-TAMRA	26	993	426
Reverse	5'-cttctggcacatgtctacac-3'	21	1019	427

Table BJB. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4184, Run 268695207	Tissue Name	Rel. Exp.(%) Ag4184, Run 268695207
Colon cancer 1	0.0	Bladder cancer NAT 2	0.0
Colon NAT 1	0.0	Bladder cancer NAT 3	0.0

Colon cancer 2	0.0	Bladder cancer NAT 4	0.0
Colon cancer NAT 2	0.0	Adenocarcinoma of the prostate 1	0.0
Colon cancer 3	0.0	Adenocarcinoma of the prostate 2	0.0
Colon cancer NAT 3	0.0	Adenocarcinoma of the prostate 3	0.0
Colon malignant cancer 4	0.0	Adenocarcinoma of the prostate 4	0.0
Colon normal adjacent tissue 4	0.0	Prostate cancer NAT 5	0.0
Lung cancer 1	0.0	Adenocarcinoma of the prostate 6	0.0
Lung NAT 1	0.0	Adenocarcinoma of the prostate 7	0.0
Lung cancer 2	0.0	Adenocarcinoma of the prostate 8	0.0
Lung NAT 2	0.0	Adenocarcinoma of the prostate 9	0.0
Squamous cell carcinoma 3	0.0	Prostate cancer NAT 10	0.0
Lung NAT 3	0.0	Kidney cancer 1	0.0
metastatic melanoma 1	0.0	Kidney NAT 1	0.0
Melanoma 2	100.0	Kidney cancer 2	0.0
Melanoma 3	0.0	Kidney NAT 2	0.0
metastatic melanoma 4	0.0	Kidney cancer 3	0.0
metastatic melanoma 5	0.0	Kidney NAT 3	0.0
Bladder cancer 1	0.0	Kidney cancer 4	0.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	0.0
Bladder cancer 2	0.0		

CNS_neurodegeneration_v1.0 Summary: Ag4184 Results from one experiment with the CG99944-01 gene are not included. The amp plot indicates that there were experimental difficulties with this run.

General_screening_panel_v1.4 Summary: Ag4184 Expression of the CG99944-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

Panel 4.1D Summary: Ag4184 Expression of the CG99944-01 gene is low/undetectable (CTs > 35) across all of the samples on this panel (data not shown).

General oncology screening panel_v_2.4 Summary: Ag4184 Highest expression of the CG99944-01 gene is detected exclusively in a melanoma sample (CT=34). Therefore, expression of this gene can be used to distinguish this sample from other samples used in this panel and therapeutic modulation of this gene product may be beneficial in the treatment of melanoma.

BK. CG99963-01: Cyclophilin 18

Expression of gene CG99963-01 was assessed using the primer-probe set Ag4160, described in Table BKA. Results of the RTQ-PCR runs are shown in Tables BKB, BKC, 5 BKD and BKE.

Table BKA. Probe Name Ag4160

Primers	Sequences	Length	Start Position	SEQ ID No
Forward	5'-ggcaagaccagcaagaagat-3'	20	545	428
Probe	TET-5'-caccattgctgactgtggacaactct-3'-TAMRA	26	565	429
Reverse	5'-aaaggaatggtctggtggtt-3'	20	617	430

Table BKB. CNS_neurodegeneration_v1.0

Tissue Name	Rel. Exp.(%) Ag4160, Run 215342806	Tissue Name	Rel. Exp.(%) Ag4160, Run 215342806
AD 1 Hippo	13.4	Control (Path) 3 Temporal Ctx	6.5
AD 2 Hippo	30.4	Control (Path) 4 Temporal Ctx	35.6
AD 3 Hippo	8.0	AD 1 Occipital Ctx	20.0
AD 4 Hippo	8.5	AD 2 Occipital Ctx (Missing)	0.0
AD 5 Hippo	84.7	AD 3 Occipital Ctx	6.0
AD 6 Hippo	37.9	AD 4 Occipital Ctx	25.5
Control 2 Hippo	64.2	AD 5 Occipital Ctx	84.7
Control 4 Hippo	7.9	AD 6 Occipital Ctx	29.3
Control (Path) 3 Hippo	8.0	Control 1 Occipital Ctx	4.4
AD 1 Temporal Ctx	13.5	Control 2 Occipital Ctx	98.6
AD 2 Temporal Ctx	40.1	Control 3 Occipital Ctx	14.9
AD 3 Temporal Ctx	5.0	Control 4 Occipital Ctx	8.1
AD 4 Temporal Ctx	23.5	Control (Path) 1 Occipital Ctx	100.0
AD 5 Inf Temporal Ctx	71.2	Control (Path) 2 Occipital Ctx	11.5
AD 5 Sup Temporal Ctx	26.1	Control (Path) 3 Occipital Ctx	3.7
AD 6 Inf Temporal Ctx	41.5	Control (Path) 4 Occipital Ctx	11.3
AD 6 Sup Temporal Ctx	39.2	Control 1 Parietal Ctx	7.0
Control 1 Temporal Ctx	5.6	Control 2 Parietal Ctx	25.0
Control 2 Temporal Ctx	65.1	Control 3 Parietal Ctx	23.3
Control 3 Temporal Ctx	18.8	Control (Path) 1 Parietal Ctx	94.0
Control 3 Temporal Ctx	6.6	Control (Path) 2 Parietal Ctx	26.2
Control (Path) 1	67.4	Control (Path) 3 Parietal Ctx	4.7

Temporal Ctx			
Control (Path) 2 Temporal Ctx	45.4	Control (Path) 4 Parietal Ctx	49.3

Table BKC. General_screening_panel_v1.4

Tissue Name	Rel. Exp.(%) Ag4160, Run 221297229	Tissue Name	Rel. Exp.(%) Ag4160, Run 221297229
Adipose	5.0	Renal ca. TK-10	25.3
Melanoma* Hs688(A).T	20.9	Bladder	13.5
Melanoma* Hs688(B).T	16.6	Gastric ca. (liver met.) NCI-N87	20.6
Melanoma* M14	31.6	Gastric ca. KATO III	75.8
Melanoma* LOXIMVI	25.5	Colon ca. SW-948	14.7
Melanoma* SK-MEL-5	52.9	Colon ca. SW480	77.4
Squamous cell carcinoma SCC-4	27.0	Colon ca.* (SW480 met) SW620	33.9
Testis Pool	2.7	Colon ca. HT29	30.8
Prostate ca.* (bone met) PC-3	25.0	Colon ca. HCT-116	66.4
Prostate Pool	3.7	Colon ca. CaCo-2	27.4
Placenta	1.7	Colon cancer tissue	14.4
Uterus Pool	2.6	Colon ca. SW1116	8.0
Ovarian ca. OVCAR-3	16.7	Colon ca. Colo-205	7.8
Ovarian ca. SK-OV-3	18.0	Colon ca. SW-48	13.2
Ovarian ca. OVCAR-4	14.1	Colon Pool	8.1
Ovarian ca. OVCAR-5	52.5	Small Intestine Pool	3.0
Ovarian ca. IGROV-1	27.4	Stomach Pool	1.6
Ovarian ca. OVCAR-8	25.5	Bone Marrow Pool	3.7
Ovary	5.8	Fetal Heart	5.5
Breast ca. MCF-7	29.5	Heart Pool	4.2
Breast ca. MDA-MB-231	60.3	Lymph Node Pool	5.1
Breast ca. BT 549	45.4	Fetal Skeletal Muscle	4.2
Breast ca. T47D	100.0	Skeletal Muscle Pool	1.9
Breast ca. MDA-N	14.5	Spleen Pool	2.6
Breast Pool	3.9	Thymus Pool	6.7
Trachea	4.8	CNS cancer (glio/astro) U87-MG	40.9
Lung	1.9	CNS cancer (glio/astro) U-118-MG	54.7
Fetal Lung	11.2	CNS cancer (neuro;met) SK-N-AS	30.4
Lung ca. NCI-N417	10.4	CNS cancer (astro) SF-539	28.7

Lung ca. LX-1	31.0	CNS cancer (astro) SNB-75	51.4
Lung ca. NCI-H146	9.0	CNS cancer (glio) SNB-19	16.2
Lung ca. SHP-77	42.3	CNS cancer (glio) SF-295	28.9
Lung ca. A549	26.8	Brain (Amygdala) Pool	7.2
Lung ca. NCI-H526	10.7	Brain (cerebellum)	10.3
Lung ca. NCI-H23	20.4	Brain (fetal)	9.2
Lung ca. NCI-H460	9.3	Brain (Hippocampus) Pool	6.7
Lung ca. HOP-62	13.8	Cerebral Cortex Pool	9.4
Lung ca. NCI-H522	22.5	Brain (Substantia nigra) Pool	7.9
Liver	1.6	Brain (Thalamus) Pool	6.8
Fetal Liver	11.8	Brain (whole)	6.3
Liver ca. HepG2	9.6	Spinal Cord Pool	3.3
Kidney Pool	7.8	Adrenal Gland	9.1
Fetal Kidney	11.3	Pituitary gland Pool	1.9
Renal ca. 786-0	20.0	Salivary Gland	3.7
Renal ca. A498	8.7	Thyroid (female)	3.8
Renal ca. ACHN	13.6	Pancreatic ca. CAPAN2	30.8
Renal ca. UO-31	21.6	Pancreas Pool	10.4

Table BKD. Panel 4.1D

Tissue Name	Rel. Exp.(%) Ag4160, Run 173118878	Tissue Name	Rel. Exp.(%) Ag4160, Run 173118878
Secondary Th1 act	57.8	HUVEC IL-1beta	70.7
Secondary Th2 act	48.3	HUVEC IFN gamma	51.4
Secondary Tr1 act	53.6	HUVEC TNF alpha + IFN gamma	31.4
Secondary Th1 rest	12.3	HUVEC TNF alpha + IL4	36.6
Secondary Th2 rest	13.0	HUVEC IL-11	30.4
Secondary Tr1 rest	14.9	Lung Microvascular EC none	58.2
Primary Th1 act	58.2	Lung Microvascular EC TNFalpha + IL-1beta	48.3
Primary Th2 act	64.2	Microvascular Dermal EC none	43.2
Primary Tr1 act	68.3	Microvascular Dermal EC TNFalpha + IL-1beta	39.8
Primary Th1 rest	25.7	Bronchial epithelium TNFalpha + IL1beta	37.4
Primary Th2 rest	14.5	Small airway epithelium none	22.8
Primary Tr1 rest	29.5	Small airway epithelium TNFalpha + IL-1beta	42.3
CD45RA CD4 lymphocyte	47.6	Coronary artery SMC rest	33.0

act			
CD45RO CD4 lymphocyte act	77.4	Coronary artery SMC TNFalpha + IL-1beta	27.4
CD8 lymphocyte act	69.7	Astrocytes rest	18.4
Secondary CD8 lymphocyte rest	53.6	Astrocytes TNFalpha + IL-1beta	17.8
Secondary CD8 lymphocyte act	24.0	KU-812 (Basophil) rest	43.8
CD4 lymphocyte none	6.7	KU-812 (Basophil) PMA/ionomycin	49.3
2ry Th1/Th2/Tr1_anti-CD95 CH11	24.8	CCD1106 (Keratinocytes) none	51.1
LAK cells rest	28.7	CCD1106 (Keratinocytes) TNFalpha + IL-1beta	49.0
LAK cells IL-2	31.2	Liver cirrhosis	7.6
LAK cells IL-2+IL-12	26.1	NCI-H292 none	71.2
LAK cells IL-2+IFN gamma	28.7	NCI-H292 IL-4	90.8
LAK cells IL-2+ IL-18	26.8	NCI-H292 IL-9	100.0
LAK cells PMA/ionomycin	38.4	NCI-H292 IL-13	85.9
NK Cells IL-2 rest	28.7	NCI-H292 IFN gamma	71.7
Two Way MLR 3 day	26.1	HPAEC none	33.0
Two Way MLR 5 day	44.4	HPAEC TNF alpha + IL-1 beta	49.3
Two Way MLR 7 day	30.4	Lung fibroblast none	29.3
PBMC rest	9.6	Lung fibroblast TNF alpha + IL-1 beta	20.4
PBMC PWM	53.2	Lung fibroblast IL-4	31.2
PBMC PHA-L	47.0	Lung fibroblast IL-9	52.5
Ramos (B cell) none	82.9	Lung fibroblast IL-13	28.7
Ramos (B cell) ionomycin	81.2	Lung fibroblast IFN gamma	43.5
B lymphocytes PWM	43.2	Dermal fibroblast CCD1070 rest	47.0
B lymphocytes CD40L and IL-4	37.6	Dermal fibroblast CCD1070 TNF alpha	62.4
EOL-1 dbcAMP	36.9	Dermal fibroblast CCD1070 IL-1 beta	33.2
EOL-1 dbcAMP PMA/ionomycin	28.5	Dermal fibroblast IFN gamma	30.6
Dendritic cells none	39.5	Dermal fibroblast IL-4	56.6
Dendritic cells LPS	27.7	Dermal Fibroblasts rest	42.3
Dendritic cells anti-CD40	26.1	Neutrophils TNFa+LPS	3.7

Monocytes rest	11.7	Neutrophils rest	5.4
Monocytes LPS	14.9	Colon	7.1
Macrophages rest	37.4	Lung	13.1
Macrophages LPS	18.3	Thymus	20.7
HUVEC none	37.9	Kidney	40.1
HUVEC starved	51.1		

Table BKE. General oncology screening panel_v_2.4

Tissue Name	Rel. Exp.(%) Ag4160, Run 268624163	Tissue Name	Rel. Exp.(%) Ag4160, Run 268624163
Colon cancer 1	32.5	Bladder cancer NAT 2	0.5
Colon cancer NAT 1	11.6	Bladder cancer NAT 3	0.3
Colon cancer 2	52.9	Bladder cancer NAT 4	6.2
Colon cancer NAT 2	16.6	Adenocarcinoma of the prostate 1	22.7
Colon cancer 3	85.9	Adenocarcinoma of the prostate 2	1.6
Colon cancer NAT 3	19.1	Adenocarcinoma of the prostate 3	9.0
Colon malignant cancer 4	100.0	Adenocarcinoma of the prostate 4	25.5
Colon normal adjacent tissue 4	10.0	Prostate cancer NAT 5	4.2
Lung cancer 1	39.2	Adenocarcinoma of the prostate 6	6.0
Lung NAT 1	0.8	Adenocarcinoma of the prostate 7	6.4
Lung cancer 2	82.4	Adenocarcinoma of the prostate 8	2.2
Lung NAT 2	1.9	Adenocarcinoma of the prostate 9	27.5
Squamous cell carcinoma 3	18.3	Prostate cancer NAT 10	0.9
Lung NAT 3	1.7	Kidney cancer 1	15.0
metastatic melanoma 1	36.6	Kidney NAT 1	5.1
Melanoma 2	5.3	Kidney cancer 2	59.0
Melanoma 3	6.7	Kidney NAT 2	15.2
metastatic melanoma 4	22.5	Kidney cancer 3	20.6
metastatic melanoma 5	72.7	Kidney NAT 3	4.9
Bladder cancer 1	0.6	Kidney cancer 4	20.0
Bladder cancer NAT 1	0.0	Kidney NAT 4	11.0
Bladder cancer 2	10.2		

CNS_neurodegeneration_v1.0 Summary: Ag4160 This panel confirms the expression of the CG99963-01 gene at low levels in the brain in an independent group of individuals. This gene is found to be slightly down-regulated in the temporal cortex of
5 Alzheimer's disease patients. Therefore, up-regulation of this gene or its protein product,

or treatment with specific agonists for this protein may be useful in reversing the neuronal death and dementia/memory loss associated with this disease.

The CG99963-01 gene codes for cyclophilin 18 (Cyclophilin A; CyP-A) homolog. Cyp-A, a soluble cytoplasmic immunophilin, is known for its involvement in T cell differentiation and proliferation. Although CyP-A has a pivotal role in the immune response, it is most highly concentrated in brain. It is known to play a role in neuronal differentiation and proliferation of human embryonic brain cells (Nahreini et al., 2001, Cell Mol Neurobiol 21(1):65-79, PMID: 11440199). Therefore, therapeutic modulation of Cyp-A like protein encoded by this gene may be useful in treatment of neurological disorders.

General_screening_panel_v1.4 Summary: Ag4160 Highest expression of the CG99963-01 gene is detected in a breast cancer T47D cell line (CT=21). High expression of this gene is also seen in cluster of cancer cell lines including melanoma, squamous cell carcinoma, pancreatic, CNS, colon, gastric, renal, breast, ovarian, and prostate cancer cell lines. Therefore, therapeutic modulation of this gene through the use of small molecule drug may be beneficial in the treatment of these cancers.

Among tissues with metabolic or endocrine function, this gene is expressed at high to moderate levels in pancreas, adipose, adrenal gland, thyroid, pituitary gland, skeletal muscle, heart, liver and the gastrointestinal tract. Therefore, therapeutic modulation of the activity of this gene may prove useful in the treatment of endocrine/metabolically related diseases, such as obesity and diabetes.

In addition, this gene is expressed at high levels in all regions of the central nervous system examined, including amygdala, hippocampus, substantia nigra, thalamus, cerebellum, cerebral cortex, and spinal cord. Therefore, this gene may play a role in central nervous system disorders such as Alzheimer's disease, Parkinson's disease, epilepsy, multiple sclerosis, schizophrenia and depression.

Panel 4.1D Summary: Ag4160 Highest expression of the CG99963-01 gene is detected in IL-9 treated NCI-H292 cells (CT=23). This gene is expressed at high to moderate levels in a wide range of cell types of significance in the immune response in health and disease. These cells include members of the T-cell, B-cell, endothelial cell, macrophage/monocyte, and peripheral blood mononuclear cell family, as well as epithelial and fibroblast cell types from lung and skin, and normal tissues represented by colon,

lung, thymus and kidney. This ubiquitous pattern of expression suggests that this gene product may be involved in homeostatic processes for these and other cell types and tissues. This pattern is in agreement with the expression profile in

General_screening_panel_v1.4 and also suggests a role for the gene product in cell

- 5 survival and proliferation. Therefore, modulation of the gene product with a functional therapeutic may lead to the alteration of functions associated with these cell types and lead to improvement of the symptoms of patients suffering from autoimmune and inflammatory diseases such as asthma, allergies, inflammatory bowel disease, lupus erythematosus, psoriasis, rheumatoid arthritis, and osteoarthritis.

- 10 **General oncology screening panel_v_2.4 Summary:** Ag4160 Highest expression of the CG99963-01 gene is detected in malignant colon cancer (CT=23). In addition, high expression of this gene is also detected in number of cancer samples including squamous cell carc

15

OTHER EMBODIMENTS

Although particular embodiments have been disclosed herein in detail, this has been done by way of example for purposes of illustration only, and is not intended to be limiting with respect to the scope of the appended claims, which follow. In particular, it is

- 5 contemplated by the inventors that various substitutions, alterations, and modifications may be made to the invention without departing from the spirit and scope of the invention as defined by the claims. The choice of nucleic acid starting material, clone of interest, or library type is believed to be a matter of routine for a person of ordinary skill in the art with knowledge of the embodiments described herein. Other aspects, advantages, and
- 10 modifications considered to be within the scope of the following claims. The claims presented are representative of the inventions disclosed herein. Other, unclaimed inventions are also contemplated. Applicants reserve the right to pursue such inventions in later claims.

What is claimed is:

What is claimed is:

1. An isolated polypeptide comprising the mature form of an amino acid sequenced selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101
2. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101.
3. An isolated polypeptide comprising an amino acid sequence which is at least 95% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101.
4. An isolated polypeptide, wherein the polypeptide comprises an amino acid sequence comprising one or more conservative substitutions in the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101.
5. The polypeptide of claim 1 wherein said polypeptide is naturally occurring.
6. A composition comprising the polypeptide of claim 1 and a carrier.
7. A kit comprising, in one or more containers, the composition of claim 6.
8. The use of a therapeutic in the manufacture of a medicament for treating a syndrome associated with a human disease, the disease selected from a pathology associated with the polypeptide of claim 1, wherein the therapeutic comprises the polypeptide of claim 1.

9. A method for determining the presence or amount of the polypeptide of claim 1 in a sample, the method comprising:
 - (a) providing said sample;
 - (b) introducing said sample to an antibody that binds immunospecifically to the polypeptide; and
 - (c) determining the presence or amount of antibody bound to said polypeptide, thereby determining the presence or amount of polypeptide in said sample.
10. A method for determining the presence of or predisposition to a disease associated with altered levels of expression of the polypeptide of claim 1 in a first mammalian subject, the method comprising:
 - a) measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and
 - b) comparing the expression of said polypeptide in the sample of step (a) to the expression of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, said disease,wherein an alteration in the level of expression of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to said disease.
11. A method of identifying an agent that binds to the polypeptide of claim 1, the method comprising:
 - (a) introducing said polypeptide to said agent; and
 - (b) determining whether said agent binds to said polypeptide.
12. The method of claim 11 wherein the agent is a cellular receptor or a downstream effector.
13. A method for identifying a potential therapeutic agent for use in treatment of a pathology; wherein the pathology is related to aberrant expression or aberrant physiological interactions of the polypeptide of claim 1, the method comprising:

- (a) providing a cell expressing the polypeptide of claim 1 and having a property or function ascribable to the polypeptide;
- (b) contacting the cell with a composition comprising a candidate substance; and
- (c) determining whether the substance alters the property or function ascribable to the polypeptide;

whereby, if an alteration observed in the presence of the substance is not observed when the cell is contacted with a composition in the absence of the substance, the substance is identified as a potential therapeutic agent.

14. A method for screening for a modulator of activity of or of latency or predisposition to a pathology associated with the polypeptide of claim 1, said method comprising:
 - (a) administering a test compound to a test animal at increased risk for a pathology associated with the polypeptide of claim 1, wherein said test animal recombinantly expresses the polypeptide of claim 1;
 - (b) measuring the activity of said polypeptide in said test animal after administering the compound of step (a); and
 - (c) comparing the activity of said polypeptide in said test animal with the activity of said polypeptide in a control animal not administered said polypeptide, wherein a change in the activity of said polypeptide in said test animal relative to said control animal indicates the test compound is a modulator activity of or latency or predisposition to, a pathology associated with the polypeptide of claim 1.
15. The method of claim 14, wherein said test animal is a recombinant test animal that expresses a test protein transgene or expresses said transgene under the control of a promoter at an increased level relative to a wild-type test animal, and wherein said promoter is not the native gene promoter of said transgene.
16. A method for modulating the activity of the polypeptide of claim 1, the method comprising contacting a cell sample expressing the polypeptide of claim 1 with a

compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.

17. A method of treating or preventing a pathology associated with the polypeptide of claim 1, the method comprising administering the polypeptide of claim 1 to a subject in which such treatment or prevention is desired in an amount sufficient to treat or prevent the pathology in the subject.
18. The method of claim 17, wherein the subject is a human.
19. A method of treating a pathological state in a mammal, the method comprising administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101 or a biologically active fragment thereof.
20. An isolated nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 101.
21. The nucleic acid molecule of claim 20, wherein the nucleic acid molecule is naturally occurring.
22. A nucleic acid molecule, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NO: 2n-1, wherein n is an integer between 1 and 101.
23. An isolated nucleic acid molecule encoding the mature form of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:2n, wherein n is an integer between 1 and 101.

24. An isolated nucleic acid molecule comprising a nucleic acid selected from the group consisting of $2n-1$, wherein n is an integer between 1 and 101.
25. The nucleic acid molecule of claim 20, wherein said nucleic acid molecule hybridizes under stringent conditions to the nucleotide sequence selected from the group consisting of SEQ ID NO: $2n-1$, wherein n is an integer between 1 and 101, or a complement of said nucleotide sequence.
26. A vector comprising the nucleic acid molecule of claim 20.
27. The vector of claim 26, further comprising a promoter operably linked to said nucleic acid molecule.
28. A cell comprising the vector of claim 26.
29. An antibody that immunospecifically binds to the polypeptide of claim 1.
30. The antibody of claim 29, wherein the antibody is a monoclonal antibody.
31. The antibody of claim 29, wherein the antibody is a humanized antibody.
32. A method for determining the presence or amount of the nucleic acid molecule of claim 20 in a sample, the method comprising:
 - (a) providing said sample;
 - (b) introducing said sample to a probe that binds to said nucleic acid molecule;and
 - (c) determining the presence or amount of said probe bound to said nucleic acid molecule,thereby determining the presence or amount of the nucleic acid molecule in said sample.
33. The method of claim 32 wherein presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type.

34. The method of claim 33 wherein the cell or tissue type is cancerous.
35. A method for determining the presence of or predisposition to a disease associated with altered levels of expression of the nucleic acid molecule of claim 20 in a first mammalian subject, the method comprising:
- a) measuring the level of expression of the nucleic acid in a sample from the first mammalian subject; and
 - b) comparing the level of expression of said nucleic acid in the sample of step (a) to the level of expression of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease;
- wherein an alteration in the level of expression of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.
36. A method of producing the polypeptide of claim 1, the method comprising culturing a cell under conditions that lead to expression of the polypeptide, wherein said cell comprises a vector comprising an isolated nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 101.
37. The method of claim 36 wherein the cell is a bacterial cell.
38. The method of claim 36 wherein the cell is an insect cell.
39. The method of claim 36 wherein the cell is a yeast cell.
40. The method of claim 36 wherein the cell is a mammalian cell.
41. A method of producing the polypeptide of claim 2, the method comprising culturing a cell under conditions that lead to expression of the polypeptide, wherein said cell comprises a vector comprising an isolated nucleic acid molecule

comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO:2n-1, wherein n is an integer between 1 and 101.

42. The method of claim 41 wherein the cell is a bacterial cell.
43. The method of claim 41 wherein the cell is an insect cell.
44. The method of claim 41 wherein the cell is a yeast cell.
45. The method of claim 41 wherein the cell is a mammalian cell.

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
6 February 2003 (06.02.2003)

PCT

(10) International Publication Number
WO 03/010327 A3(51) International Patent Classification⁷: C12P 21/06,
C12N 1/20, 15/00, C07K 1/00, C07H 21/04

(21) International Application Number: PCT/US02/14199

(22) International Filing Date: 2 May 2002 (02.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/288,063	2 May 2001 (02.05.2001)	US
60/288,395	3 May 2001 (03.05.2001)	US
60/289,087	7 May 2001 (07.05.2001)	US
60/289,818	9 May 2001 (09.05.2001)	US
60/289,817	9 May 2001 (09.05.2001)	US
60/290,194	11 May 2001 (11.05.2001)	US
60/290,753	14 May 2001 (14.05.2001)	US
60/291,181	15 May 2001 (15.05.2001)	US
60/291,243	16 May 2001 (16.05.2001)	US
60/292,001	18 May 2001 (18.05.2001)	US
60/292,374	21 May 2001 (21.05.2001)	US
60/292,587	22 May 2001 (22.05.2001)	US
60/293,107	23 May 2001 (23.05.2001)	US
60/293,747	25 May 2001 (25.05.2001)	US
60/294,110	29 May 2001 (29.05.2001)	US
60/294,109	29 May 2001 (29.05.2001)	US
60/294,434	30 May 2001 (30.05.2001)	US
60/294,827	31 May 2001 (31.05.2001)	US
60/304,879	12 July 2001 (12.07.2001)	US
60/308,901	31 July 2001 (31.07.2001)	US
60/312,270	14 August 2001 (14.08.2001)	US
60/313,416	17 August 2001 (17.08.2001)	US
60/318,463	10 September 2001 (10.09.2001)	US
60/325,683	27 September 2001 (27.09.2001)	US
60/330,292	18 October 2001 (18.10.2001)	US
60/333,873	28 November 2001 (28.11.2001)	US
60/336,909	3 December 2001 (03.12.2001)	US
60/337,552	3 December 2001 (03.12.2001)	US
60/359,245	21 February 2002 (21.02.2002)	US
10/136,826	1 May 2002 (01.05.2002)	US

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:

US	60/288,395 (CIP)
Filed on	3 May 2001 (03.05.2001)

US	60/308,901 (CIP)
Filed on	31 July 2001 (31.07.2001)
US	60/330,292 (CIP)
Filed on	18 October 2001 (18.10.2001)
US	60/288,063 (CIP)
Filed on	2 May 2001 (02.05.2001)
US	60/289,087 (CIP)
Filed on	7 May 2001 (07.05.2001)
US	60/289,818 (CIP)
Filed on	9 May 2001 (09.05.2001)
US	60/313,416 (CIP)
Filed on	17 August 2001 (17.08.2001)
US	60/289,817 (CIP)
Filed on	5 September 2001 (05.09.2001)
US	60/290,194 (CIP)
Filed on	11 May 2001 (11.05.2001)
US	60/325,683 (CIP)
Filed on	27 September 2001 (27.09.2001)
US	60/336,909 (CIP)
Filed on	3 December 2001 (03.12.2001)
US	60/290,753 (CIP)
Filed on	14 May 2001 (14.05.2001)
US	60/291,181 (CIP)
Filed on	15 May 2001 (15.05.2001)
US	60/291,243 (CIP)
Filed on	16 May 2001 (16.05.2001)
US	60/292,001 (CIP)
Filed on	18 May 2001 (18.05.2001)
US	60/292,374 (CIP)
Filed on	21 May 2001 (21.05.2001)
US	60/312,270 (CIP)
Filed on	14 August 2001 (14.08.2001)
US	60/292,587 (CIP)
Filed on	22 May 2001 (22.05.2001)
US	60/359,245 (CIP)
Filed on	21 February 2001 (21.02.2001)
US	60/293,107 (CIP)
Filed on	23 May 2001 (23.05.2001)
US	60/294,110 (CIP)
Filed on	29 May 2001 (29.05.2001)
US	60/304,879 (CIP)
Filed on	12 July 2001 (12.07.2001)
US	60/293,747 (CIP)
Filed on	25 May 2001 (25.05.2001)

[Continued on next page]

(54) Title: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

(57) Abstract: The present invention provides novel isolated polynucleotides and small molecule target polypeptides encoded by the polynucleotides. Antibodies that immunospecifically bind to a novel small molecule target polypeptide or any derivative, variant, mutant or fragment of that polypeptide, polynucleotide or antibody are disclosed, as are methods in which the small molecule target polypeptide, polynucleotide and antibody are utilized in the detection and treatment of a broad range of pathological states. More specifically, the present invention discloses methods of using recombinantly expressed and/or endogenously expressed proteins in various screening procedures for the purpose of identifying therapeutic antibodies and therapeutic small molecules associated with diseases.



WO 03/010327 A3



US 60/294,434 (CIP)
 Filed on 30 May 2001 (30.05.2001)
 US 60/318,463 (CIP)
 Filed on 10 September 2001 (10.09.2001)
 US 60/294,109 (CIP)
 Filed on 29 May 2001 (29.05.2001)
 US 60/333,873 (CIP)
 Filed on 28 November 2001 (28.11.2001)
 US 60/337,552 (CIP)
 Filed on 3 December 2001 (03.12.2001)
 US 60/294,827 (CIP)
 Filed on 31 May 2001 (31.05.2001)
 US Not furnished (CIP)
 Filed on 1 May 2002 (01.05.2002)

(71) **Applicant** (for all designated States except US): **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th floor, New Haven, CT 06511 (US).

(72) **Inventors; and**

(75) **Inventors/Applicants** (for US only): **MILLER, Charles, E.** [US/US]; 98 Saddle Hill Drive, Guilford, CT 06437 (US). **KEKUDA, Ramesh** [IN/US]; 71 Aiken Street, Unit R3, Norwalk, CT 06851 (US). **MALYANKAR, Uriel, M.** [IN/US]; 229 Branford Road, Number 330, Branford, CT 06405 (US). **LI, Li** [CN/US]; 56 Jerimoth Drive, Branford, CT 06405 (US). **PENA, Carol, E., A.** [US/US]; 604 Orange Street, Number 2, New Haven, CT 06511 (US). **SPYTEK, Kimberly, A.** [US/US]; 7 Hotel Road, Ellington, CT 06029 (US). **GORMAN, Linda** [US/US]; 329 Monticello Drive, Branford, CT 06405 (US). **GUO, Xiaojia** [CN/US]; 713 Robert Frost Drive, Branford, CT 06405 (US). **FERNANDES, Elma, R.** [IN/US]; 77 Florence Road, 2B, Branford, CT 06405 (US). **SMITHSON, Glenda** [US/US]; 125 Michael Drive, Guilford, CT 06405 (US). **STONE, David, J.** [US/US]; 223 Whitehorn Drive, Guilford, CT 06437 (US). **ZERHUSEN, Bryan, D.** [US/US]; 337 Monticello Drive, Branford, CT 06405 (US). **PATTURAJAN, Meera** [IN/US]; 45 Harrison Avenue, Apartment 1C, Branford, CT 06405 (US). **ANDERSON, David, W.** [US/US]; 59 Hacienda Circle, Plantsville, CT 06479 (US). **MEZES, Peter, S.** [CA/US]; 7 Clark's Lane, Old Lyme, CT 06371 (US). **PEYMAN, John, A.** [US/US]; 336 West Rock Avenue, New Haven, CT 06515 (US). **MACDOUGALL, John, R.** [CA/US]; 117 Russell Street, Hamden, CT 06517 (US). **PADIGARU, Muralidhara** [IN/US]; 71 Hampton Park, Branford, CT 06405 (US).

RASTELLI, Luca [IT/US]; 52 Pepperbush Lane, Guilford, CT 06437 (US). **SHENOY, Suresh, G.** [IN/US]; 15 Millwood Drive, Branford, CT 06405 (US). **GERLACH, Valerie, L.** [US/US]; 18 Rock Pasture Road, Branford, CT 06405 (US). **SHIMKETS, Richard, A.** [US/US]; 5 Indian Meadows Drive, Guilford, CT 06437 (US). **ZHONG, Mei** [CA/US]; 152 Brushy Plain Road, Branford, CN 06405 (US). **EDINGER, Shlomit, R.** [US/US]; 766 Edgewood Avenue, New Haven, CT 06515 (US). **ELLERMAN, Karen** [US/US]; 87 Montoya Drive, Branford, CT 06405 (US).

(74) **Agent:** **ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris, Glovsky and Popeo, P. C., One Financial Center, Boston, MA 02111 (US).

(81) **Designated States (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(88) **Date of publication of the international search report:**
 24 December 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/14199

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12P 21/06; C12N 1/20, 15/00; C07K 1/00; C07H 21/04
US CL : 435/69.1, 252.3, 320.1, 530/350, 536/23.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/69.1, 252.3, 320.1, 530/350, 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
BIOSIS, CAPLUS, EMBASE, MEDLINE, SCISEARCH, USPTO WEST, GENBANK

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	Database GenBank, US National Library of Medicine, (Bethesda, MD, USA) No. AC025904, BIRREN et al., May 2000.	20-28, 36-45
X	CHOU, C.F. et al. Identification of a keratin-associated protein that localizes to a membrane compartment. Biochem. J. January 1994, Vol. 298, No. 2, pages 457-463, entire document.	1-7
X	LIAO, J. et al. Association of glucose-regulated protein (grp78) with human keratin 8. FEBS Lett. January 1997, Vol. 417, No. 3, pages 316-320, entire document.	1-7



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

16 October 2003 (16.10.2003)

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

Facsimile No. (703)305-3230

Date of filing of the international search report

Authorized officer

Manjunath N. Rao

Telephone No. 703-306-0196

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/14199

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-7, 20-28, 36-45, SEQ ID NO:1 and 2 only

Remark on Protest

☐
☐

The additional search fees were accompanied by the applicant's protest.

No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US02/14199

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group 1 through Group 101, claims 1-7, 20-28, 36-45, drawn to polynucleotides and polypeptides comprising nucleotide sequences and amino acid sequences which are chemically unrelated, do not require each other for practice and have separate utilities, vectors comprising polynucleotides, host cells and method of making the polypeptide wherein each group consists of a single polynucleotide and its corresponding single polypeptide starting from SEQ ID NO:1 to SEQ ID NO:101 (i.e., each group comprises one DNA and its respective encoded polypeptide sequence).

Group 102, claim(s) 8, drawn to use of a polypeptide as a therapeutic.

Group 103, claim(s) 9, drawn to a method of determining the presence or amount of the polypeptide.

Group 104, claim(s) 10, drawn to a method of determining the presence of a disease.

Group 105, claim(s) 11-12, drawn to method of identifying an agent that binds to the polypeptide.

Group 106, claim(s) 13, drawn to a method of identifying an agent for use in treatment of a pathology.

Group 107, claim(s) 14-15, drawn to a method of screening for a modulator of a polypeptide.

Group 108, claim(s) 16, drawn to a method of modulating the activity of the polypeptide.

Group 109, claim(s) 17-18, drawn to a method of preventing or treating a pathology associated with a polypeptide.

Group 110, claim(s) 19, drawn to a method of treating a pathological state in a mammal by using the polypeptide.

Group 111, claim(s) 29-31, drawn to an antibody.

Group 112, claim(s) 32-34, drawn to a method of determining the presence of nucleic acid.

Group 113, claim(s) 35, drawn to a method of determining the presence of or predisposition to a disease.

The inventions listed as Groups 1-212 do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Pursuant to 37 C.F.R. 1.475 (d), the ISA/US considers that where multiple products and processes are claimed, the main invention shall consist of the first invention of the category first mentioned in the claims and the first recited invention of each of the other categories related thereto. Accordingly, the main invention (Group I) comprises the first-recited product, a polynucleotide encoding its corresponding polypeptide, a vector, a host cell, a method for producing polypeptide. Further pursuant to 37 C.F.R. 1.475 (d), the ISA/US considers that any feature which the subsequently recited products and methods share with the main invention does not constitute a special technical feature within the meaning of PCT Rule 13.2 and that each of such products and methods accordingly defines a separate invention.

